UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA DE LISBOA

Reinforcement learning across development in humans

Rita Alexandra Figueira Belo

Dissertação orientada pelo Professor
Doutor Tiago Vaz Maia

Mestrado em Neurociências

2015

Todas as afirmações efectuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Medicina de Lisboa pelos conteúdos nele apresentados.

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 22 de Setembro de 2015.

"If one day you have to choose between the world and love, remember:

If you choose the world you'll be left without love,

but **if you choose love, with it you will conquer the world**."

Albert Einstein

Acknowlegments/Agradecimentos

Apesar deste ser, na maioria dos casos, um gesto demasiado pequeno em comparação com o apoio e amizade recebidos por todos vós, gostaria de deixar aqui o meu agradecimento às pessoas sem as quais este trabalho não teria sido possível.

Em primeiro lugar, gostaria de agradecer ao meu orientador, ao Professor Tiago Maia, por me ter possibilitado desenvolver este projecto autonomamente, acreditando na minha intuição e nas minhas capacidades. Obrigada também pela compreensão e simpatia nos momentos mais difíceis.

Também gostaria de agradecer a todos os meus colegas de laboratório, em especial ao Ângelo Dias, à Catarina Farinha, e ao Vasco Conceição pelos bons momentos, pela amizade e pela preocupação. Ao Ângelo, o meu companheiro de projecto, obrigada por tudo. Foste e serás sempre um excelente professor. Um sincero obrigada tanto ao Vasco como ao Tiago Henriques pela ajuda e partilha de conhecimentos na fase final do trabalho. Agradeço também à Ana Portêlo e à Lena Ernst pelos seus bons conselhos e ajuda na revisão final do documento.

À comunidade escolar do Agrupamento de Escolas João de Barros, deixo o meu profundo agradecimento. Um especial obrigada aos directores da Escola Secundária João de Barros, o Professor António Carvalho, da Escola 2/3 ciclos de Corroios, a Professora Sílvia Rodrigues, e da Escola EB1/JI do Miratejo, a Professora Adelina Ramos, por possibilitarem que as actividades inerentes a este projecto fossem desenvolvidas nas referidas instituições. Agradeço também ao Professor Paulo Silva, o empréstimo dos computadores portáteis utilizados nas suas aulas, bem como toda a ajuda para que os mesmos estivessem prontos para as actividades. Sem esses computadores, nunca teria conseguido atingir o nível de participações a que cheguei. Um especial obrigado ao Professor Manuel Jorge que, desde o dia em que lhe bati à porta, tornou-se a pessoa a quem mais tenho a agradecer. Muito obrigada pelo seu profissionalismo, empenho e dedicação a este projecto, assim como pela sua amizade e boa-disposição.

Também não poderia deixar de agradecer ao Centro Cultural e Recreativo do Alto do Moinho (CCRAM), em especial à Ana Cristina Videira, cuja sua amizade e estima possibilitaram não só a divulgação deste projecto junto dos utilizadores do CCRAM, mas também a sua consequente participação no mesmo.

À Alexandra Botelho, obrigada pela ajuda (burocrática ou não) e incentivo durante esta etapa. A tua devoção aos estudantes é exemplar!

Agradeço também às minhas amigas Carolina Toste e Natália Madeira pelos momentos de descontração, pelas parvoíces e pelos desabafos! À Isa Silvestre, um exemplo a seguir, um sincero obrigada por me teres ajudado a acreditar novamente em mim. Um obrigada também às minhas queridas amigas Mariana Oliveira, Liliana Brito e Diana Marques, cuja amizade valorizo muito. Obrigada por estarem sempre lá, por me ouvirem e, mais importante que tudo, pelo vosso carinho. Estão todas num cantinho especial…

Como não poderia deixar de ser, um GRANDE obrigada ao pessoal do Departamento de Informática da maravilhosa FCT-UNL (vocês sabem todos quem são), e em especial ao André Grossinho, pela animada companhia durante as longas horas de trabalho, pela distração e pelo espírito académico durante este último ano de trabalho. Ao Ricardo Monteiro e ao António Lopez, muito obrigada pela revisão do texto.

Agradeço também ao meu maravilhoso namorado, Gonçalo Tavares, por teres conseguido aturar toda a minha essência ao longo destes últimos (já 9) anos. Obrigada por sonhares comigo, e por me trazeres de volta à Terra.

À minha irmã agradeço aqueles momentos de cumplicidade de manas (apesar de serem cada vez menos)... Obrigada por me deixares orgulhosa em tantos momentos.

Por fim, gostaria de agradecer aos meus pais a possibilidade de estudar com as melhores condições possíveis – significou sempre muito para mim.

Ao meu pai, muito obrigado por me teres ajudado a transportar aquilo que pareciam ser infinitos computadores portáteis e pela companhia naqueles dias difíceis de recolha de dados. Tornaste os meus dias menos solitários.

À minha mãe, a minha melhor amiga, obrigada por sempre me compreenderes, pelo incentivo e palavras de força. Obrigada por todo esse amor incondicional que nem sempre sei como retribuir. Este trabalho é tanto meu como teu porque sem ti nunca teria chegado onde cheguei hoje. Obrigada por tudo.

Por último, um grande obrigada a todos os participantes e respectivos encarregados de educação (no caso dos mais novinhos), pelo interesse neste projecto. Sem vocês não haveria projecto!

Para o meu avô Figueira

Resumo

Hoje em dia, o tratamento de muitas doenças neuropsiquiátricas foca-se na atenuação dos sintomas exibidos pelos doentes, não havendo, na maioria dos casos, um conhecimento profundo sobre as disfunções neurológicas que se encontram na origem das mesmas, nem sobre os mecanismos de ação pelos quais a medicação exerce os seus efeitos. Como tal, nos últimos anos tem-se verificado um esforço colectivo, nomeadamente no campo das neurociências computacionais, para a construção de novas ferramentas que permitam avaliar quantitivamente determinadas funções cognitivas e assim obter informação importante sobre a patofisiologia das doenças acima referidas.

Espera-se que estas novas ferramentas tenham particular impacto ao nível das patologias do neurodesenvolvimento, devido à elevada comorbilidade existente entre estas, bem como às semelhanças entre as manifestações sintomáticas exibidas pelos diferentes doentes. Devido ao foco nestas doenças, um dos circuitos mais estudados é o circuito cortico-estriado-talamo-cortical, visto que este se encontra frequentemente significativamente comprometido nos doentes que sofrem destas patologias, tendo este circuito um papel fulcral no processo de tomadas de decisão e na motivação. Sabe-se também que a dopamina tem um papel modelador fundamental para o normal funcionamento deste circuito, nomeadamente devido à sua atividade neuromoduladora durante a aprendizagem por reforços.

No estriado, os níveis de dopamina modulam a actividade das vias directa (ou *Go*) e indirecta (ou *NoGo*) envolvidas no processo de aprendizagem por reforços. Estas vias são antagónicas, sendo que (de um modo simplista) quando uma via se encontra activa a outra está inibida. Assim, a presença de dopamina no estriado leva à activação e inibição das vias directa e indirecta, respectivamente, sendo que a sua ausência provoca um efeito contrário.

No processo de aprendizagem por reforços, as acções executadas pelo agente são reforçadas ou inibidas, de acordo com o valor atribuído não só ao reforço recebido, mas também ao valor das previsões feitas. Neste sentido, a actividade dopaminérgica no estriado codifica os erros de previsão obtidos

durante o processo de aprendizagem por reforços. Quando os erros de previsão são positivos, ou seja, o valor do reforço (neste caso, da recompensa) é maior do que o esperado *a priori*, existe um aumento da actividade dopaminérgica basal no estriado, levando à posterior activação da via directa (e inibição da via indirecta). Pelo contrário, quando os erros são negativos, o valor do reforço (neste caso, da punição) é inferior ao esperado, o que diminui a actividade basal dopaminérgica e leva à activação da via indirecta (e inibição da via directa.

Desta forma, ao longo dos últimos anos têm sido desenvolvidas várias tarefas cognitivas computorizadas para o estudo do processo de aprendizagem por reforços, tanto em pessoas saudáveis como em doentes. Estes estudos têm contribuído, por um lado, para uma maior compreensão do processo de aprendizagem por reforços e, por outro, ajudado a melhor compreender a fisiopatologia de determinadas doenças neuropsiquiátricas.

Neste contexto, o presente estudo teve como principal objectivo determinar a influência da idade no processo de aprendizagem por reforços em humanos. Para tal, entre Novembro de 2014 e Fevereiro de 2015, foram recrutados participantes saudáveis com idades compreendidas entre os 6 e os 80 anos. Para serem admitidos no estudo, todos os participantes cumpriram a série de critérios seguidamente apresentada: gestação não inferior a 36 semanas; sem historial de 1) doenças neurológicas e psiquiátricas, 2) traumatismos cranianos com perda de consciência, e 3) ataques epiléticos; ausência de medicação psiquiátrica e/ou neurológica nos últimos 4 meses. Além disso, todos os participantes tinham o português como língua materna. Posto isto, para avaliar a influência da idade na aprendizagem por reforços e nas tendências motoras, foi aplicada uma nova tarefa probabilística de aprendizagem por reforços, bem como um questionário final a todos os participantes.

A tarefa aplicada é semelhante a um jogo de computador simples. Nesta, são apresentadas aleatoriamente 5 imagens (30 ensaios para cada uma) e o participante, dependendo da acção executada (carregar ou não na barra de espaços do teclado) pode receber pontos (+1), perder pontos (-1) ou nem ganhar nem perder pontos (0). Desta forma, o objectivo dos participantes é tentar obter o maior número de pontos possíveis sendo que, para tal, estes têm de aprender qual a melhor acção a desempenhar para cada imagem. Cada jogo tem 5 condições aleatoriamente atribuídas para cada imagem: *Go to win* (carregar para ganhar pontos), *NoGo to win* (não carregar para ganhar pontos), *Go to avoid losing* (carregar para evitar perder pontos), *NoGo to avoid losing* (não carregar para evitar perder pontos), e Neutra (na qual, carregar ou não dá sempre zero pontos). Exceptuando na condição Neutra, o ganho ou perda de pontos tem uma probabilidade associada. Na condição Neutra, o resultado associado a uma das acções é fixo (100% hipótese de receber zero pontos), e nas restantes segue uma relação de 80% (+1 ou -1) / 10% (-1 ou +1) / 10% (0).

Por sua vez, o questionário final foi dividido em duas partes com o objectivo de, na primeira, recolher informações demográficas acerca dos participantes (por exemplo: idade, sexo, ano de escolaridade), e, na segunda, recolher informações que permitam avaliar o nível de consciência dos participantes em relação ao seu próprio desempenho durante a execução da tarefa.

Neste estudo, tanto a tarefa como o questionário foram aplicados a 419 pessoas saudáveis (cuja média de idades é de 17.36 ± 10.78 anos, e dos quais 51.6% são sexo masculino), distribuídos por 18 faixas etárias desde os 6 aos 80 anos. É no entanto de salientar que 92.4% dos participantes tinham idade inferior a 30 anos.

A análise dos resultados comportamentais permitiu concluir que de facto a idade dos sujeitos condiciona fortemente a aprendizagem por reforços, mas não tem muita influência nas tendências motoras. Através da utilização desta tarefa foi também possível identificar uma ordem temporal específica de aprendizagem das condições analisadas, uma vez que a capacidade de aprender um maior número de condições é influenciada pela idade. De um modo geral, a partir dos 9 anos de idade, os participantes foram capazes de aprender ambas as condições congruentes (*Go to win* e *NoGo to avoid losing*), nas quais a acção é "compatível" com o valor intrínseco da condição. Entre os 11-13 anos de idade, observou-se o início da aprendizagem da condição *Go to avoid losing*, sendo que, só por volta dos 15 anos de idade é que os sujeitos se começaram a mostrar capazes de aprender a condição *NoGo to win*. Estas últimas são consideradas condições incongruentes, uma vez que a acção favorável é contrária ao valor intrínseco da condição.

De forma a entender com maior exactidão qual a relação entre o desempenho da tarefa, dado pelas proporções de respostas correctas dos últimos 20 ensaios de cada condição, e as idades dos participantes, procedeu-se ao desenvolvimento de um modelo matemático. Nesta análise foram incluídos os resultados comportamentais dos participantes com idades compreendidas entre os 6 e os 30 anos, uma vez que não foram recolhidos dados suficientes para crer que os resultados obtidos para os participantes mais velhos possam ser representativos. Desta forma, a partir deste modelo matemático, foi possível determinar a existência de uma relação logarítmica entre o desempenho da tarefa e a idade dos participantes (*p-value* $= 1.18\ e^{-94}$).

O modelo mostrou-se capaz de prever a sequência temporal de aprendizagem das diferentes condições supracitada. Adicionalmente, a utilização de contrastes permitiu comparar os diferentes valores assimptóticos de aprendizagem de cada condição, confirmando uma maior aprendizagem entre as condições congruentes, quando comparadas com as incongruentes (p-value ~ 0), e entre as condições cuja acção favorável era carregar na tecla de espaço, *Go to win* e *Go to avoid losing*, e entre aquelas que era melhor não carregar, *NoGo to win* e *NoGo to avoid losing* (p-value = 0.008). Não se verificaram diferenças significativamente estatísticas entre as condições maioritariamente associadas a uma aprendizagem por estímulos positivos (condições *win*), e por estímulos negativos (condições *avoid losing*).

Por fim, o questionário permitiu verificar que os sujeitos demontraram alguma percepção acerca do seu desempenho quando questionados após a execução da tarefa, uma vez que, em média, os sujeitos conseguiram identificar correctamente a melhor acção a executar para cada condição. Em relação a quais os pontos que costumavam aparecer mais vezes no ecrã para cada condição, os sujeitos demonstraram alguma dificuldade o que pode traduzir a falta de confiança nas suas respostas. Por outro lado, quando foi pedido uma classificação quanto ao gosto pela estética de dada imagem, inconscientemente os participantes atribuíram melhor classificação às condições *win*, e pior classificação às condições *avoid losing* e *Neutra.*

Apesar de algumas limitações, este estudo oferece um avanço para uma melhor compreensão do processo de aprendizagem por reforços em humanos,

em particular sobre a influência da idade no mesmo. Esperamos também que, num futuro próximo, esta informação adquira uma ainda maior importância, ao permitir que os mecanismos específicos associados à etiologia e patofisiologia das doenças neurológicas e psiquiátricas possam ser estudados de uma forma mais controlada e exacta, tanto através desta como através de abordagens neurocomputacionais semelhantes.

Palavras-chave

Aprendizagem por reforços; Idade; Tarefa probabilística Go/NoGo; Tendências motoras

Abstract

The etiology of most neuropsychiatric disorders is currently unknown and treatment strategies concerning these disorders are mainly targeted at the amelioration of symptoms. To overcome this, several approaches aiming to study not only distinct cognitive processes but also how these processes can be disrupted in the aforementioned disorders have been developed in recent years.

Neurodevelopment disorders, in particular, constitute a significant burden to our society. For this reason, we developed a new probabilistic task to evaluate reinforcement learning (RL) and motor biases in children, adolescents, and young adults, since these aspects are related to basal-ganglia functioning and dopaminergic signalling, two processes which are commonly reported to be impaired in these patients.

To this end, we used our task and a final questionnaire. Our sample included 419 healthy subjects, aged from 6 to 80 years old (mean = 17.36 ± 10.78 years; 51.6% males), most of whom (92.4%) were less than 30 years old.

We found that RL performance increased with age, and that subjects were mostly unaware of their performance. In fact, our results identified a sequence across age by which the subjects learned the task contingencies. From 9 years old, subjects were capable of learning both to win points and to avoid losing points by, respectively, pressing or withholding from pressing a key, during task solving (congruent learning). Around 11-13 years old, the subjects started to learn to avoid losing points by pressing the key, and only at 15 years old, they started to learn to win points by not pressing the key (incongruent learning). A general linear model of task performance across age also predicted the aforementioned sequence ($p\text{-}value = 1.18\ e^{-94}$). Through this model, we found once again that subjects were better in congruent than in incongruent learning ($p\text{-}value \sim 0$), and that their performance was better in conditions where the correct action was to perform a key press ($p\text{-}value = 0.008$).

Key-words: Age; Motor biases; Probabilistic Go/NoGo task; Reinforcement learning

Abbreviations list

ADHD	Attention deficit and hyperactivity disorder
Adj-R^2	Adjusted-squared-R^2
AIC	Akaike information criterion
ANOVA	Analysis of variance
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
BG	Basal-ganglia
BIC	Bayesian information criterion
CA	Correct action
cAMP	Cyclic adenosine monophosphate
CCRAM	*Centro Cultural e Recreativo do Alto do Moinho*
CEV	Constant expected value
CEVR	Constant expected value - reverse
CN	Caudate nucleus
CR	Conditioned response
CS	Conditioned stimulus
CSTC	Cortico-striatal-thalamo-cortical
DA	Dopamine
DEV	Decreased expected value
DRD1	Dopamine receptor D1
DRD2	Dopamine receptor D2
EB1/JI	*Escola Básica 1º ciclo e Jardim de Infância do Miratejo*
EB23C	*Escola Básica 2, 3 de Corroios*
ESJB	*Escola Secundária João de Barros*
EV	Expected value
fMRI	Functional magnetic resonance imaging
Gpe	Globus pallidum extern
Gpi	Globus pallidum intern
G-proteins	Guanosine nucleotide-binding proteins

IEV	Increased expected value
KDE	Kernel density estimation
NA	Nucleus accumbens
NMDA	*N*-methyl-D-aspartate
O	Outcome
OCD	Obsessive-compulsive disorder
PD	Parkinson's disease
PFC	Pre-frontal cortex
PKA	Protein kinase A
PLC	Phospholipase C
Pt	Putamen
R	Response
R^2	Squared-R
RL	Reinforcement learning
S	Stimulus
SD	Standard deviation
SNc	Substantia nigra pars compacta
SNr	Substantia nigra pars reticulata
STN	Subthalamic nucleus
Str	Striatum
TS	Tourette syndrome
UR	Unconditioned response
US	Unconditioned stimulus
VTA	Ventral tegmental area

Index

Index of Figures

Index of Tables

Introduction

In this chapter, the main theoretical aspects of this thesis are briefly introduced. It includes some psychology, neuroscience and computer science concepts, as well as the interaction between them in the context of this work.

1. Learning

Learning can be defined as "*the process by which changes in behaviour arise as a result of experience interacting with the world*" (Gluck, Mercado, & Myers, 2008). This definition highlights the fact that the act of learning produces knowledge, or memories, which may lead to behaviour modifications.

In an evolutionary perspective, species became so well-adapted to their natural environment through their evolutionary process of natural selection, which was defined for the first time by Charles Darwin (1809-1882) (Darwin, 2009). In fact, for thousands of years, species had the possibility of learning how to survive by adapting their behaviour to certain situations, so that the best behaviour was kept through natural selection. The necessary knowledge to make the best decision, from a range of available options, became crucial to organisms' survival. For instance, a rat that does not have fear of cats, does not possess the knowledge that cats are the enemy, will certainly die in a cat confrontation. Therefore, the organisms more capable of learning and with better memory have more survival hypothesis and pass these traits to their offspring (Gluck, Mercado, & Myers, 2008)

In short, learning is the process used by organisms to acquire knowledge about the world, while memory is the process by which this knowledge is encoded, stored and later retrieved (Kandel, Schwartz, & Jessell, 2000). Furthermore, memories are "*the record of our past experiences acquired through learning*" (Gluck, Mercado, & Myers, 2008).

The basic mechanism of memory formation has been highly conserved over billions of years of biological evolution. This mechanism requires a "plastic brain" that can be physically modified by life experiences (Kandel, Schwartz, & Jessell, 2000). There are more than one memory systems, and since the majority of these concepts were defined during the last decades, there are still several disagreements among scientific community (Squire, 2004; Gluck, Mercado, & Myers, 2008).

In the light of this work, it will only be considered the distinction between 1) explicit memory, which includes semantic and episodic memory; and 2) implicit memory, which includes procedure memory and habituation, among others (Squire, 2004). This distinction depends on whether or not these memories were consciously and intentionally recollected, respectively. Briefly, the explicit memory is known to be hippocampal-dependent, and was largely perceived by the famous case of the Henry Molaison (better known as H.M. patient), who had both hippocampus partially removed in an attempt to cure his severe and refractory temporal lobe epilepsy (Scoville & Milner, 1957). On the other hand, the implicit memory depends mainly on the basal-ganglia activity, which will be explained later.

Moreover, to learn about something, it is crucial to recognize signals, which may be stimuli to act. First of all, animals should not react to all exterior signals, since some of these may not represent anything to them. Thus, after learning the value of each environmental signal through habit learning, animals can filter them, where the habitual neutral signals begin to be ignored, so that the relevant ones are easily detected (Gleitman, 1995). This topic will be better described in section *1.2. Habit learning.*

Nowadays, one of the most studied topics in behavioural neuroscience is the decision-making process for selecting actions, which is also the main topic of this work. For the last two centuries, the decision-making process has been extensively studied by using classic and instrumental conditioning methods. More recently, an interesting symbiosis between biological neuroscience and artificial intelligence concepts has enabled some new strategies for studying these emerged mechanisms, leading to progress in both biological and computational fields.

During this section, some of these topics will be briefly explored, in order to help understanding the underling work of this study.

1.1. Conditioning

The associations between stimulus and action may be studied by conditioning, leading to the behavioural adaption where some responses become more frequent than others. In short, conditioning allows “learning to anticipate a positive event and preparing to take maximal advantage of it” as well as “to anticipate negative events” (Gluck, Mercado, & Myers, 2008).

Moreover, conditioning can be divided into two types: classical and instrumental. Very briefly, in classical conditioning, a stimulus (S) becomes associated to a response (R), so subjects must learn the S-R association. In instrumental learning, the response is learned through the given outcome (O), so subjects learn the R-O association (Gluck, Mercado, & Myers, 2008).

1.1.1. Classical conditioning

In late 19^{th} Century, Ivan P. Pavlov (1849-1936) changed the behavioural thinking of that time with his works in animal conditioning (Pavlov, 1927). Although Pavlov’s works started with the study of digestion, which had already given him a Nobel Prize in 1904, his greatest contribution was accomplished when he started studying the digestive reflexes in dogs, mainly the salivation. In fact, his findings became known as classical conditioning, or pavlovian conditioning (Gleitman, 1995).

Through his work, Pavlov found that dog’s salivation began to be triggered by a cluster of stimuli that initially were completely neutral. At the beginning, only the food in the dog’s mouth caused salivation but after a while in the laboratory, dogs started to salivate just by looking at food, looking at the food’s empty plate or at the person that usually brought the food, or even by listening his steps.

To study this process, Pavlov used a basic procedure which includes two types of stimuli: a neutral or conditioned stimulus (CS), and a rewarding or unconditioned stimulus (US). The CS must not elicit a specific response and, on contrary, the US must be capable of eliciting a specific response, which is also called an unconditioned response (UR). The aim of this procedure was that, at some point, a CS began to elicit a conditioned response (CR), which should be quite similar to the UR. To better understand this experience, it is essential to clearly distinguish conditioned trials from unconditioned trials. As the name suggests, conditioned trials are used for conditioning, which imply that the CS is paired with the US, so that the CR is reinforced. On the other hand, in unconditioned trials the US is not presented so that the CR is not reinforced. Therefore, the CR begins to appear after a period of conditioning, where the conditioned trials are presented.

Figure 1 illustrates the three steps of this procedure: a) before conditioning, b) conditioning, and c) after conditioning, where some tasty dog's food and a tone were used as the US and CS, respectively. At first (figure 1, A.), only the food (US) elicited the dog's salivation - the dog's UR to food. During conditioning (figure 1, B.), which consists in pairing several times the tone (CS) with the food (US), the dog eventually started to salivate when it heard the tone. After a while, the dog's salivation appears just by hearing the tone, even if the food was not presented (figure 1, C.), so this became a conditioned response (CR).

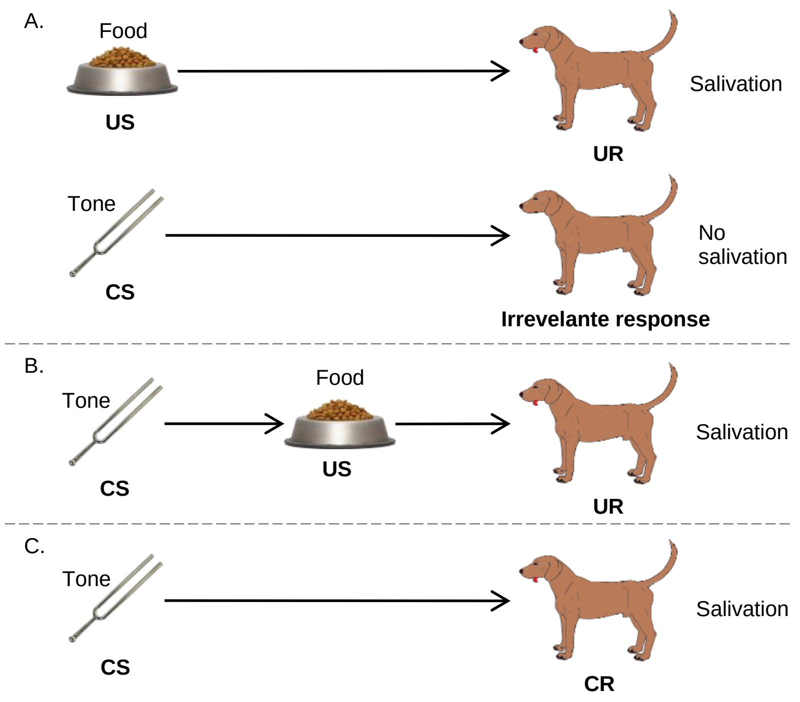


Figure 1. Pavlov's classical conditioning procedure. A. Before conditioning; B. During conditioning; C. After conditioning. US – unconditioned stimulus; UR – unconditioned response; CS – conditioned stimulus; CR – conditioned response.

Classical conditioning has some features which must be considered. First of all, the acquisition of the CR is directly dependent of the proportion between conditioned and unconditioned trials: while the first reinforces the CR, the second does not. Therefore, these trials are also commonly called reinforced and non-reinforced trials, respectively.

Moreover, the CR's strength increases with the number of conditioned trials and can be measured in several ways: one can be through the amplitude of the response in unconditioned trials (e.g. the amount of dog's saliva in Pavlov's experiments); it can be measured by the probability of response, which is given by the proportion of response trials in unconditioned trials; and through the latency between CS's presentation and the response's start. In opposition to the first two measures that increased with the CR's strength, the latency decreases with CR's strength. Furthermore, an established CR can be gradually extinct if no more conditioning trials are presented, so that CS stops representing the reward, being the CR not further reinforced. However, a stimulus that was already extinguished can be reconditioned through a new conditioning phase. This reconditioning is usually faster than if no associations were previously made, since it needs less conditioned trials for that CR to recover the same strength.

Finally, the stimuli can be generalized and discriminated, two important aspects during organisms' evolution. The generalization allows that organisms express the same behaviour in presence of similar stimuli. This is essential in organisms' survival because, in nature, stimuli are not exactly the same and might suffer some changes, although they still imply the same learned response. On the other hand, discrimination is needed when stimuli are identically but elicit contrary behaviours, so their generalization may be catastrophic to organisms. In both cases, the similarity between stimuli will affect the organisms' capability of generalizing and discriminating them: the more similar the stimuli are, the easier their generalization process and the harder their discrimination process, and *vice-versa* (Gleitman, 1995).

1.1.2. Instrumental conditioning

Instrumental conditioning, also known as instrumental learning, differs from classical conditioning essentially in the kind of conditional response. In instrumental conditioning, the reward essentially depends on organism's response, which is chosen from a cluster of available responses. Thus, animals need to establish a response-outcome (R-O) association, instead of a stimulus-response (S-R) association, as in classical conditioning. Summarily, in classical conditioning, animals learn the temporal association between CS-US, which is independent from the animals' response, so the CR is forced by the US; while in instrumental conditioning, animals learn the relation between their responses and the reward, wherein the best response maximizes the reward (Gleitman, 1995).

In 1898, Eduard L. Thorndike (1874-1949) presented a method for studying instrumental learning, which consisted in presenting problem for animals to solve (Thorndike, 1898). He put starving cats inside a "puzzle box" where, in order to exit, cats needed to learn to do a specific procedure, like pull a string. When cats were well succeeded, they received a small food portion.

At the beginning of these experiments, cats tried to leave the box by all means, such as scratching and biting the box's bars unsuccessfully. After a while, they also started to look for a way to escape inside the box and they were eventually successful. Considering this, Thorndike described that cats performed progressively better over time when trying to solve the box problem, instead of starting to perform perfectly at some point during the experiment, which was expected if cats *understood* the correct answer in that precise moment. Therefore, cats' performance was successively improved with the number of trials, measured by the time spent by cats to come out of the box – or the response latency.

Later, Thorndike's achievements were postulated as the famous *Law of Effect* (Thorndike, 1911). According to this law, during the learning process some responses are reinforced while others are not, thus the outcome of each response is essential in this process. If one response leads to the desirable situation, then this response will be reinforced by its own outcome – that is the desirable situation. On the other hand, if other response does not lead to the desirable situation, or even if it leads to a punishment, that response will be weakened. After a while, the animal's tendency to respond will change according to the strength of each experienced response.

This law corroborated the evolutionary thinking of the time, namely Darwin's theory of evolution: the adaptive nature of animals' behaviour largely depends on their biological needs (Alexander, 1974). The *Law of Effect* is similar to the *Law of survival of the fittest*, where genetic well-adapted animals have better chances of surviving and thus transmitting their genetic pool to their offspring. In animals' life, the *Law of Effect* determines that the animal's best responses will survive (Alexander, 1974).

B. F. Skinner (1904-1990) clearly distinguished classic from instrumental conditioning: in classic conditioning, he claimed that animal behaviour is evoked by the CS, which is an environmental stimulus; and in instrumental conditioning the behaviour "came" from inside of the animal, since it is a voluntary response, so it suffers less influence of the environment. He named these voluntaries responses *operants*, since they operate in environment to produce a reward (Skinner, 1938).

Skinner also adapted the Thorndike's method, by constructing an experimental chamber where a rat could pull a lever or a pigeon could pick a lighting square, repetitively, so that the same response could be repeated several times. This chamber, known as the Skinner's box, allowed that all responses, stimuli and rewards to be automatically given to the animal. In this case, the response strength is measured as the response rate, which is the amount of responses per time unit. The results of these responses are commonly called outcomes, feedback or reinforcers.

Until now, instrumental conditioning has been treated as a way of learning responses that lead to good situations, but this process is also a way of learning responses that avoid or minimize bad situations. In fact, there are two kinds of reinforcers: the positive reinforcer, or a reward; and the negative reinforcer, or a punishment. Aversive stimuli are commonly used as a punishment after animal's response. In this situation, the negative reinforcer weakens the response so that the tendency of the animal's response will decrease over time.

Moreover, negative stimuli might also positively reinforce responses, namely responses of escape or avoidance, since these can be a way that the animal found to stop the discomfort, leading the tendency of these responses to increase across trials. Therefore, the tendency for the correct response increases with the amount of reinforcers received for that specific response, and the tendency to a specific response might be extinguished if reinforcers stop at some time.

In addition, once that the instrumental response is not evocated by environmental stimuli, as in classical conditioning, the stimuli are used to discriminate the responses (Gleitman, 1995). For instance, a green light might indicate that if a pigeon picks the target in the next moment, it will receive a reward, but a red light might indicate that even if a pigeon picks the target, it will not receive any reward. In summary, these discriminative stimuli add information about the environment where animals operate. Similarly, the generalization of stimuli will be easier depending on the similarity between stimuli (Gleitman, 1995).

Until now, the learning process that is focused on the relation between CS-US happens only in classical conditioning. Edward C. Tolman (1886-1959) applied the same interpretation to instrumental conditioning, as he believed that animals are able to create an intern representation of the relation between response and its reinforcement (Tolman, 1932; Tolman, 1948). For instance, a rat does not learn only to press a lever but also learn that, by pressing the lever, it will receive a pellet of food. Therefore, animals learn the representation between R-O, which may be used in future situations.

1.1.3. The surprise’s effect

In classic conditioning, the CS works as a sign to the US essentially due to their contiguity. Robert Rescorla (1967 -) went further by claiming that classical conditioning is dependent not only on the contiguity of CS-US, but also on CS's absence, which is associated to the US's absence (Rescorla, 1967). Thus, the relation between US-CS is contiguous and contingent. Despite that, this contingent may be incomplete (e.g. an US may occur with 80% probability when CS is presented) and, even so, the conditioning will happen.

Other important aspect in classical conditioning is the surprising effect of US. If the US is unexpected, animals can associate it with the CS (Gleitman, 1995). Moreover, the bigger the surprise of US, the stronger the association between CS-US. Likewise, if animals are not surprised with the US, no association between CS-US will occur.

Furthermore, Leon Kamin (1927 -) developed several experiments to study this surprise's effect, which led him to discover the *Blocking Effect* in 1969 (Kamin, 1969). Kamin demonstrated that a redundant stimulus, for which a given animal already has information, will not be associated to a CS. This means that, without the surprising element of CS, the animal knows what will happen next, and thus the conditioning between the two stimuli will fail.

In 1972, this phenomenon was mathematically formulated by Robert Rescorla (1940 -) and Allan Wagner (1934 -), which was nominated as the Rescorla-Wagner model (Rescorla & Wagner, 1972):

$$\Delta V = k(\lambda - V) \ . \quad (1)$$

Considering V the expected value (or the associative CS value), and λ the unexpected value (or the US value) on a given trial. Thus, the $(\lambda - V)$ expression represents the difference between the unexpected outcome and the expected outcome. In short, it represents the level of surprise of the US when it is preceded by the CS. Therefore, at the beginning, the CS value is usually zero, since it is a neutral stimulus, so the discrepancy between the CS and the US values is very high. After a while, both CS and US values become equal, leading to the $(\lambda - V)$ expression to become zero, and consequently minimizing the level of surprise: the animal fully predicts that the CS will be followed by the US.

1.2. Habit learning

In 1974, the term *habit learning* was used for the first time by Hirsh to describe a particular type of learning process that does not use the hippocampus (Hirsh, 1974). Further, he also noticed that habit learning has similarities to the S-R learning process, since both are insensitive to contextual information.

In fact, hippocampus is responsible for the episodic memory encoding, which can be defined as a cluster of a lifetime experienced memories. Additionally, hippocampus does not only encode a full description of certain situations, but it also encodes its contextual information, such as spacio-temporal and emotional data, as its individual components. Moreover, this stored contextual information, known as flexible memories, which can be individually retrieved at any time, and are formed through a rapid encoding process (one-trial learning) (Maren & Holt, 2000; Gluck, Mercado, & Myers, 2008; Kandel, Schwartz, & Jessell, 2000).

In opposition, several studies had gradually completed the Hirsh's habit learning definition by identifying some contrary features between habitual and hippocampal-learning (Mishkin, Malamut, & Bachevalier, 1984). First of all, habit learning is a slow encoding process, requiring a larger number of trials to learn the association, where the associations formed are inflexible, since they do not possess contextual information. Later, habit learning was also considered an unconscious or implicit learning with an automatic processing. It seems to be an evolutionarily early process so it is similar across species (Seger & Spiering, 2011; Redgrave, Prescott, & Gurney, 1999).

Furthermore, in the 1980s, Dickinson proposed a separation between *habit instrumental* (S-R) and *goal-direct behaviour* (R-O) learning systems (Dickinson, 1985). This division was based on whether a performed behaviour depends or not on the reward value, and these conclusions were achieved by reinforcer devaluation studies (Adams & Dickinson, 1981; Colwill & Rescorla, 1985). For example, if the animal is in food deprivation, and the reward is a small food pellet, the reinforcer devaluation will be to feed the animal before the experiments. Thus, in the case that the animal learns the action by habit, the action is a devaluation-insensitive behaviour, which means that it will still perform the action to obtain food. In humans, this kind of behaviour happens, for instance, when there is an electrical failure and we continue to turn the light switches on and off. On the other hand, in a goal-direct behaviour, the reward devaluation compromises the execution of the action, because the goal (the food) was already achieved (Yin & Knowlton, 2006; Dayan & Niv, 2008).

Additionally, habit learning can also be defined as the tendency to stop responding to certain stimuli that become habitual (Gleitman, 1995). As mentioned previously, this biological process was an important evolutionary step, allowing organisms to recognize *real* danger stimuli, which are essential to trigger fight or flight reactions in order to survive. Thus, this capability became essential to filter habitual stimuli, which did not represent a real danger to them from others. The association between a stimulus and its value can be learned through conditioning process.

2. Decision-making process

Decision-making is a cognitive process of selecting a specific option among a cluster of possible options, where each is expected to produce a different outcome (Lee, 2013). This present work comprises the decision-making process in humans of selecting actions through its given outcomes.

Nowadays, learning is recognized as a key player in the decision-making process (Frank, 2011). This fact leads to the mathematical formulation of several theories, being the reinforcement learning (RL) theory one of them (Sutton & Barto, 1998). In addition, from a neurobiological point of view, the researchers began to be quite interested in understanding how the human brain computes these decision-making and RL models (Cohen & Frank, 2009; Maia, 2009).

Furthermore, the developed neurocomputational models are important to the study of several psychiatric and neurological disorders, e.g. neurodevelopment disorders and Parkinson's disease (PD), since these have shown to be related to a maladaptive and aberrant decision-making process (Cohen & Frank, 2009). This symbiosis between neurobiology findings and computational models has contributed to a fully understanding of the mechanisms behind these disorders, leading to a more accurate diagnosis and efficient treatments in the future (Maia & Frank, 2011).

2.1. Model-free and model-based processes

Decision-making can be executed as a model-free and/or a model-based process (Dayan, & Niv, 2008).

To better understand these concepts we can imagine a scenario (figure 2) where a person (the agent) wants to go from one place to another. If the agent knows the path extremely well, like in cases where a person goes back home after work (represented by the dashed line - figure 2), he does the entire journey by taking the necessary decisions by habit, so “without thinking too much”. On the other hand, if the agent wants to visit a new florist to buy flowers for a friend (represented by the solid line - figure 2), he will mentally plan his journey before starting it.

In the first situation, the path for going back home from work was repeated so many times by the agent that it has become usual, so the agent selects the actions by model-free. In the second situation, the agent is moved by his goal so his behaviour is goal-direct. In this case, the agent has a mental map of the city and he chooses the best path based on that model, so his decisions are performed as a model-based behaviour.

Additionally, several studies have suggested that both strategies, model-free and model-based, are implemented in parallel by humans’ brain and, depending on the circumstances, usually with one process standing out more than the other (Dayan, & Niv, 2008).

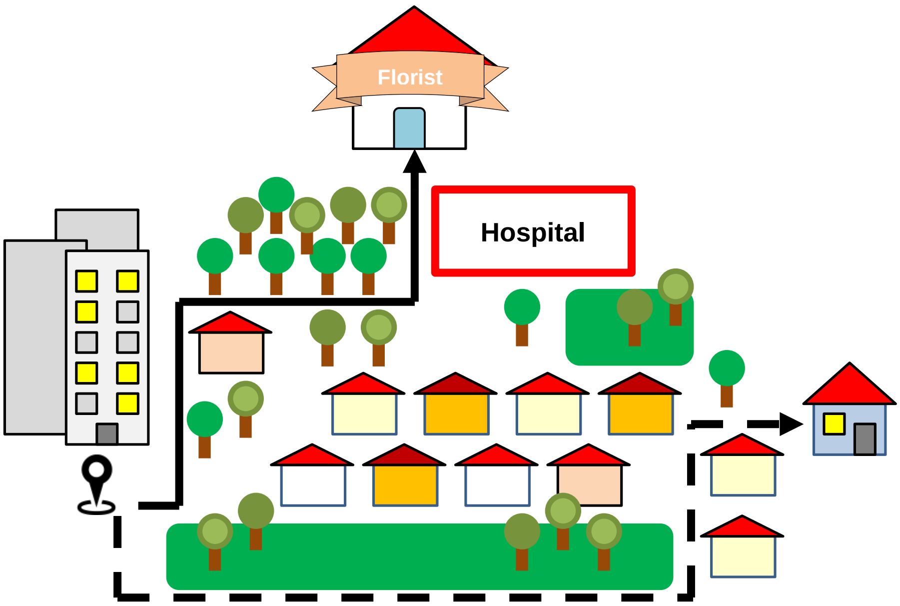


Figure 2. Representation of both model-free and model-based decision-making process within an imaginary scenario. The model-free decision-making situation: backing home by habit (dashed line); the model-based decision-making situation: visiting a new florist by planning the journey before living work (solid line).

2.2. Reinforcement learning

The reinforcement learning theory is a powerful framework to study the process of decision-making, which originally emerged from psychological theories of learning in animals. Besides being considered a branch of artificial intelligent, this theory is largely studied nowadays in order to better understand more complex daily-life choices situations (Lee, 2013; Sutton & Barto, 1998).

In RL, the decision-maker, or agent, must learn the best action among a pool of alternative actions in order to maximize rewards and diminishing punishments (Sutton & Barto, 1998). More importantly, by experience, this process often requires changes in the decision-making strategies of the agent, where he must adapt his behaviour across time.

Figure 3 represents the interaction between the agent and the surrounding environment across time. During the learning process, at each moment (t, being $t = 1,2,\ldots$), of a given state (s_t, being $s_t \in S$, and S the set of all possible states in time t), the agent must perform an action (a_t, being $a_t \in A(s_t)$, and $A(s_t)$ the cluster of all possible actions in a given stage s_t), according to a policy $\pi_t(s,a)$. The policy can be interpreted as the necessary adjustments in the agent's behaviour so that he can reach his goal.

The RL methods, namely the temporal difference models, similar to the Rescorla-Wagner model presented previously – equation (1), are able to continuously update the values (V) of each action through a reward estimation.

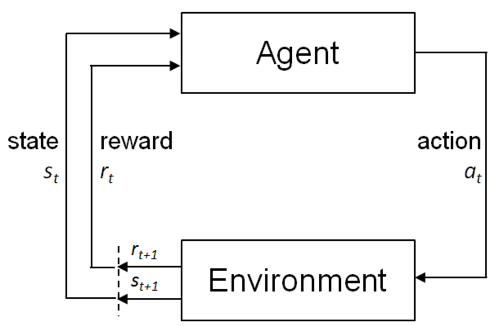


Figure 3. Agent-environment interaction in reinforcement learning (Sutton & Barto, 1998).

Therefore, these estimated values (V_{t+1}) are computed by using the prediction error (δ) of each action in a given state, which is the difference between its obtained reward (r) and the expected reward (V_t), previously estimated (Sutton & Barto, 1998).

The following equations explain this process, with α being the learning rate that dictates the update degree of the prediction error estimation in the estimated value (Sutton & Barto, 1998):

$$\delta = r - V_t \ , \tag{2}$$

$$V_{t+1} = V_t + \alpha . \delta \ . \tag{3}$$

2.3. Basal-ganglia circuitry

Mishkin and his colleagues were the first to propose that basal ganglia structures were involved in habit learning (Mishkin, Malamut, & Bachevalier, 1984), and later, its role in RL process was also recognized (Frank & Claus, 2006; Ludvig, Bellemare, & Pearson, 2011).

The basal-ganglia (BG), or basal-nuclei circuit (figure 4), is constituted by several anatomical and functionally linked subcortical nuclei, which are located at the base of the forebrain (Kandel, Schwartz, & Jessell, 2000). It includes the striatum (Str), the globus pallidum extern (GPe), and intern (GPi), the substantia nigra pars compacta (SNc), pars reticulata (SNr), and the subthalamic nucleus (STN).

The striatum itself includes the nucleus accumbens (NA), the caudate nucleus (CN), the putamen (Pt), being the ventral striatum constituted by the CN and the Pt (Kandel, Schwartz, & Jessell, 2000; Ludvig, Bellemare, & Pearson, 2011).

Moreover, the BG structures are involved in motor, associative and limbic loops, due to their strong interconnection with surrounding brain areas (Seger & Spiering, 2011; Frank, 2011; Cohen & Frank, 2009).

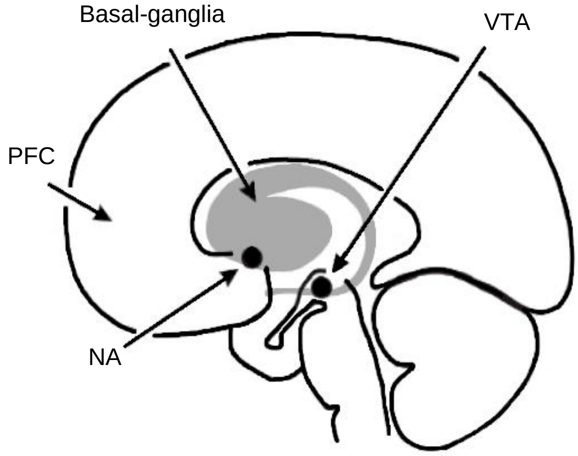


Figure 4. Representation of a sagittal section of a human brain, wherein the basal-ganglia structure is represented by the grey area. PFC – pre-frontal cortex; NA – nucleus accumbens; VTA – ventral tegmental area. Adapted from Ludvig, Bellemare, & Pearson, 2011**.**

In the context of this work, it will be given more emphasis to motor loop, namely the cortico-striatal-thalamo-cortical (CSTC) circuitry, whose neuronal activity is modulated by dopamine (DA), serotonin and acetylcholine (Kandel, Schwartz, & Jessell, 2000). Besides to the BG structures, the CSTC circuitry, as the name suggests, also includes the thalamus, and some cortical regions, namely the pre-frontal cortex (PFC), the motor cortex, the pre-motor cortex, and the sensorimotor cortex (Frank, 2011).

A schematic representation of the BG circuit (and model) is presented in figure 5, with emphasis to the dopaminergic modulation in the ventral striatum (Cohen & Frank, 2009). In fact, the striatum has an important mediating role within the brain, since it receives inputs from its many different areas, including from other structures within the CSTC circuitry, namely from the PFC, the SNc, and the thalamus. It is also considered the primary afferent structure of the CSTC circuitry, sending outputs only to structures within the circuitry. In opposition, the thalamus is the CSTC's output structure, sending output back to the cortex, namely to the motor system (Frank, 2005; Cohen & Frank, 2009).

It is broadly known that the BG has the dynamic function of selecting which actions' representations in frontal cortex may pass throughout the circuit to the motor system, leading to the execution of the motor action initially represented. In fact, BG works like a gate, facilitating and suppressing specific actions' representations in frontal cortex (Frank, Loughry, & O'Reilly, 2001; Reynolds, Hyland, & Wickens, 2001).

This process occurs essentially through the balance between the two striatal pathways: the direct and indirect pathway, where its dopaminergic modulation becomes essential (Nicola, Surmeier, & Malenka, 2000; Frank & Hutchison, 2006).

The direct pathway, or *Go* pathway, includes gabaergic striatonigral neurons (or *Go* neurons), which mainly express the excitatory dopamine receptor D1 (DRD1 or D1+), leading to the direct inhibition of the GPi/SNr areas (Albin, Young, & Penney, 1989). On the other hand, the indirect pathway, or *NoGo* pathway, includes gabaergic striatopallidal neurons, which mainly express the inhibitory dopamine receptor D2 (DRD2 or D2-), leading to the direct inhibition of the GPe (Albin, Young, & Penney, 1989). Both GPe and GPi are inhibitory structures tonically active through the STN glutamatergic inputs. Consequently, the thalamus is also tonically inhibited by the activity of the GPi (Chevalier & Deniau, 1990).

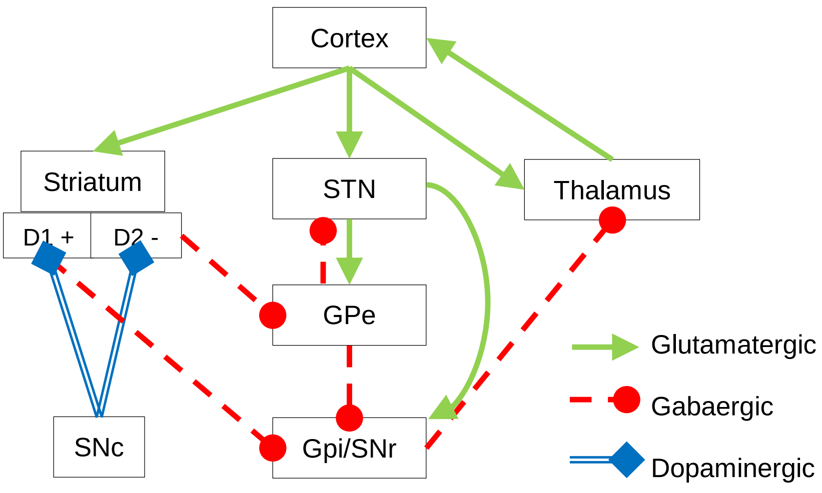


Figure 5. Diagram of the basal-ganglia model: the cortico-striatal-thalamo-cortical circuitry. STN - subthalamic nucleus; GPe – globus pallidum extern; GPi – globus pallidum intern; SNr – substantia nigra pars reticulata; SNc – substantia nigra pars compacta; D1+ – excitatory D1 receptors of dopamine; D2- – inhibitory D2 receptors of dopamine. (Adapted from Maia & Frank, 2011)

Therefore, these two pathways are complementary: when the direct pathway is active, the GPi is inhibited so the thalamus is disinhibited, leading to the potentiation of the action representation involved. On the other hand, if the indirect pathway is active, the inhibition of the GPe occurs, leading to the disinhibition of the GPi and, consequently, to the inhibition of the thalamus, thus the action representation is weakened (Cohen & Frank, 2009).

Moreover, the frontal cortex sends excitatory inputs not only to the striatum, but also to both thalamus and STN (Nambu *et al.*, 2000; Nambu, Tokuno, & Takada, 2002). In fact, the BG does not potentiate or suppress actions if they are not already represented in the thalamus (Chevalier & Deniau, 1990). In addition, the STN activation is essential for selecting the correct action representation, since STN activates both GPe and GPi, sending a global *NoGo* signal to thalamus, bypassing the striatum, until the choices are made by the latter (Bogacz & Gurney, 2007; Frank *et al.*, 2007, a). Due to this fact, the pathway between STN and GPi is also known as the *hyperdirect* pathway (Nambu, Tokuno, & Takada, 2002). Thus, this “brake” action of the STN prevents the BG to make premature choices. In fact, the amount of activity in frontal cortex is directly proportional to the sent *NoGo* signal, so the presence of simultaneous competing options increases the STN *NoGo* signal, giving more time for the striatum set the best option within all the possibilities (Frank *et al.*, 2007, a).

In short, the direct and the indirect pathways are responsible for sending a *Go* or a *NoGo* signal to the cortex, respectively; while the presence or absence of dopamine in striatum dictates the following pathways activation and suppression (Nicola, Surmeier, & Malenka, 2000). To better understand this biological process, a summary of the key aspects of the dopaminergic modulation in post-synaptic neurons is following presented.

2.3.1. Dopaminergic modulation at cellular level in brain

In the brain, dopamine is synthesized by dopaminergic neurons, mostly from mesenchephalon, which project to other brain regions. The three main dopaminergic pathways are: mesostriatal, mesolimbic, and mesocortical pathways (Nemoda, Szekely, & Sasvari-Szekely, 2011).

In this work, only the mesostriatal pathway will be considered (represented in figure 5 by the dopaminergic input in striatum – the double lines), where the dopaminergic neurons from SNc project to the ventral striatum. This pathway is essentially responsible for motor control (Joshua, Adler, & Bergman, 2009) and rewarding processes (Ikemoto & Panksepp, 1999). Moreover, dopamine acts through dopamine receptors, which belong to the seven-transmembrane receptor family, coupled to guanosine nucleotide-binding proteins (G-proteins) (Pawlak & Kerr, 2008).

The D1-like receptors, which include the dopamine receptors D1 and the D5, are associated to the $G_s\alpha$-protein. Thus they are considered excitatory receptors, leading to the activation of adenylate cyclise, consequently increasing of the cyclic adenosine monophosphate (cAMP), which activates the protein kinase A (PKA). Moreover, the PKA is responsible for the activation of the voltage-dependent calcium channel (Cav1.2/1.3), for inhibiting the voltage-dependent sodium channel (Nav1.1), and also for activating other targets responsible for the trafficking of both *N*-methyl-D-aspartate (NMDA) and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) glutamatergic receptors (Surmeier, Plotkin, & Shen, 2009).

On the other hand, the D2-like receptors, which include the dopamine receptors D2, D3, and the D4 are associated to $G_i\alpha$-protein, which is responsible for the inhibition of the adenylate cyclise, decreasing the levels of cAMP in cytosol and inactivating the PKA. In addition, $G_i\alpha$-protein also inhibits voltage-dependent calcium channel (Cav2.1/2.2), and activates the phospholipase C (PLC), which activates the voltage-dependent sodium channel (Nav1.1), and inhibits both AMPA and NMDA receptors trafficking (Surmeier, Plotkin, & Shen, 2009).

In short, dopamine signalling modulates short- and long-term glutamatergic activity in post-synaptic neurons through the activation or deactivation of both D1-type and D2-type receptors (Pawlak & Kerr, 2008; Surmeier, Plotkin, & Shen, 2009).

2.3.2. Dopaminergic signalling encodes the prediction errors

As previously mentioned, multiple studies have shown that midbrain dopaminergic activity is implicated in reward-dependent learning. In fact, the activity (or firing) of dopaminergic neurons from SNc and VTA seems to encode not only the reward value of a non-predicted stimulus, but also the prediction-error between a received reward of a predicted stimulus and its expected reward – see equation (2) (Schultz, 1998).

Figure 6 shows the pattern of dopaminergic neurons firing from SNc during a classical conditioning experiment with monkeys. These results were obtained through an extra-cellular recording of single-neurons, when the animals were performing a Go/NoGo paradigm.

Initially, several appetitive stimuli were presented to monkeys with the dopaminergic firing occurring when the animals tasted them (figure 6, A). Then, after a conditioning phase, where a pairing between a neutral stimulus (e.g. a light) and a reward occurred, the dopaminergic firing changed from the moment when the animals received the reward to the moment when these were able to predict the reward (figure 6, B). Interestingly, when the animals did not receive the expected reward, the dopaminergic neurons stopped firing briefly at the moment when they expected to receive the reward (figure 6, C) (Schultz, 1998).

These transient changes within the activity of the SNc dopaminergic neurons are responsible for phasic changes in dopamine input in the striatum. They are also known as bursts or dips of dopamine, depending on if they increase or decrease the tonic level of dopamine in the striatum (Schultz, 1998; Cachope, R., & Cheer, 2014).

Therefore, the dopamine bursts in the striatum activate both DRD1 and DRD2, leading to the activation of the direct pathway and inactivation of the indirect pathway, respectively. In opposition, a dip in striatal dopamine release impairs the activation of both dopaminergic receptors, leading to the activation of the indirect pathway and inactivation of the direct pathway (Surmeier, Plotkin, & Shen, 2009).

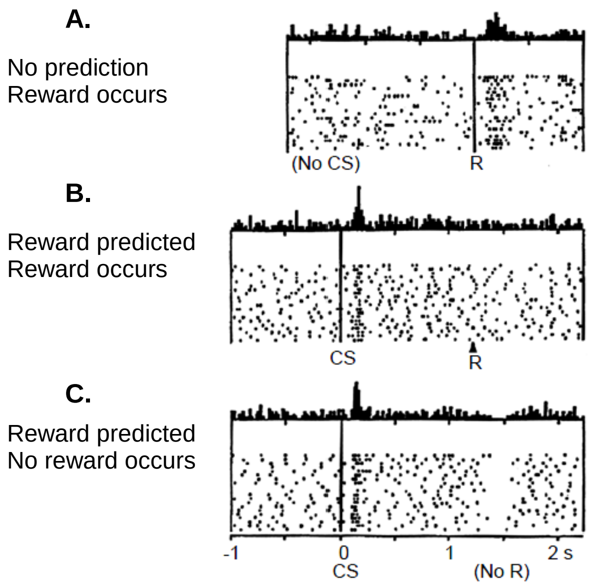


Figure 6. Pattern of dopaminergic neurons firing from SNc, during a classical conditioning experiment. A. Before conditioning; B. During conditioning; C. After conditioning. CS – conditioned stimulus; R – reward. (Schultz, 1998)

Identically, in a RL context, the feedback received after performing a specific action may produce a prediction error, leading to a phasic dopaminergic change in the striatum (Cachope, R., & Cheer, 2014). When the prediction error is positive, a burst of dopamine occurs, while a dopamine dip occurs when the prediction is negative. These two phenomena are associated to the *Go* and *NoGo* learning of that action, respectively. Thus, when a feedback is already expected, neither prediction errors nor learning process occur (Gluck, Mercado, & Myers, 2008; Bayer & Glimcher, 2005).

In addition, the amplitude of positive prediction errors is positively correlated with the dopaminergic firing rate – the burst (Bayer & Glimcher, 2005). On the other hand, the amplitude of negative prediction errors cannot be codified only by the decrease of the neurons firing rate, since its baseline is low (around 5 Hz). In fact, the duration of the pause of dopaminergic firing rate – the dips, seems to code the amplitude of negative prediction errors (Bayer, Lau, & Glimcher, 2007). Moreover, the DRD2 have greater affinity to dopamine when compared with the DRD1 (Richfield, Penney, & Young, 1989), so the duration of the dopamine dips is essential to the removal of synaptic available dopamine, leading to the inactivation of the formers and consequent activation of the *NoGo* pathway (Frank & Claus, 2006).

Other interesting aspect of striatal dopaminergic modulation is the influence of the tonic dopamine level. In general, higher levels of tonic dopamine in striatum increase the tendency of the *Go* pathway activation, and the opposite increases the tendency of the activation of the *NoGo* pathway. Thus, the tonic dopamine affects the overall tendency for motor responses, also known as the motor biases, influencing also the RL process (Maia & Frank, 2011). To explore a little bit what has been done to study RL in human, the next section includes some of these studies.

3. Reinforcement learning studies

During the last decades, the reinforcement learning process has been studied using several RL paradigms (or cognitive tasks), which have become more suitable to monitor the RL process in humans.

These tasks are usually similar to simple computerized games, where determined actions may be reinforced by positive and/or negative feedback. In addition, it is also well-know that gamification of these tasks increase people's commitment to them by making the testing moments more enjoyable for the participants (Deterding *et al.*, 2011).

The following two sections were divided into two aspects, where the first one presents some examples of previous RL studies, which highlights the greatest applicability of these tasks. Moreover, the second section briefly presents some studies aimed to understanding the age influence across the RL process in humans.

3.1. Several RL cognitive tasks and their application

These studies contributed not only to a better understanding of the basal-ganglia circuitries and the dopaminergic signalling in healthy people, but also to increase the knowledge of how the RL process might be disrupted in some neurological and psychiatric disorders, such as PD (Frank *et al*., 2007, b) and Tourette syndrome (TS) (Worbe *et al.*, 2011), respectively.

Aforementioned, the RL cognitive tasks have been designed to compass several aspects within the RL process. Having this in mind, some of these studies also included functional magnetic resonance imaging (fMRI) and genetic analyses to complement the obtained behavioural data by using these tasks.

For instance, the *temporal utility integrative* task (Moustafa *et al.*, 2008) was used to verify if time could be used as a response condition. Therefore, the design of this task included a clock face with a single course that was able to make a full turn in 5 seconds. In each trial, subjects needed to make some temporal adjustments to their single motor action, which was pressing a key, in order to maximize the reward. The task conditions were classified regarding their expected values (EV), which were computed by using the $V = probability\ x\ magnitude$ expression. Therefore, the conditions were denominated by IEV, DEV, and CEV, depending on if the EV, respectively increased, decreased, or were constant over trial. The CEVR condition was the opposite of the CEV condition: when the reward-magnitude decreased, the probability of receiving points increased over trial. In addition, the different conditions were distinguished by 4 indicative colours on the face of the clock. Therefore, the *Go* learning was represented by the speeding up required to maximize the reward in the IEV condition and, on contrary, the *NoGo* learning was represented by the act of slowing down required in the DEV condition.

This task was tested, for the first time, in patients with PD (Moustafa *et al.*, 2008). While off medication, patients (DA depleted) were better in slowing down so they showed a better performance in the DEV condition. On the other hand, if they were on medication, they were better at speeding up so they presented a better performance in the IEV condition. These results were already expected since this disorder is characterized by the striatonigral degeneration, which reduces both tonic and phasic dopaminergic levels in striatum (Frank, 2005; Frank *et al.*, 2007, a).

In contrast, different results were obtained when this task was applied to patients with schizophrenia (Strauss *et al.*, 2010), since they showed an impairment in the *Go* learning, which was more pronounced in patients with higher levels of negative symptoms, and an intact *NoGo* learning. In addition, since the neural basis of the negative symptoms remains unclear, these results suggested that a dysfunction in the *Go* pathway might be present.

Another emergent RL task used during the last decade was the Go/NoGo task developed by Frank and his colleagues (Frank, Seeberger, & O'Reilly, 2004). This task included some probabilistic Go/NoGo selection situations, which will be described below, and it was divided into a training phase and a testing phase. During the first phase, subjects needed to select one stimulus from a displayed pair of stimuli, being the feedback associated to a specific probability of winning and losing points (e.g. A: 80/20 *versus* B: 20/80). The task included three different pairs of stimuli, which were randomly displayed several times, one at a time, so that subjects might learned to choose the stimulus with the higher probability of winning points. The testing phase included new combinations of paired-stimulus, which can be divided into two categories: choose A and avoid B. As the names suggest, in the first situation choose the stimulus A was the best action, while in the second one, the best action was selecting any other stimulus to avoid selecting the B.

This task was also applied for the first time to patients with PD (Frank, Seeberger, & O'Reilly, 2004), being its results also identical to the ones obtained from the *temporal utility integrative* task: 1) patients off medications were better in avoiding B trials than in choosing A, and 2) the medication inverted this scenario, so that patients on medication were better to choose A trials. In summary, patients off medication, who had reduced levels of striatal dopamine, were more sensitive to negative feedback, showing a better *NoGo* learning than patients that had the dopamine levels restored by medication, who were more sensitive to positive feedback, showing a better *Go* learning.

Furthermore, other study designed to study the effect of some dopaminergic polymorphisms in RL process in humans also used this probabilistic Go/NoGo selection task (Frank *et al*., 2007, b). The chosen polymorphisms aimed to track down the dopaminergic influence in PFC (rs4680, or Val^{158}Met polymorphism of catechol-O-methyltransferase (COMT) gene), and the striatal *Go* pathway (rs907094 polymorphism of the dopamine- and cAMP-regulated phosphoprotein, 32 kDa (DARPP-32) gene) and *NoGo* pathway (rs6277, or C957T polymorphism of the DRD2 gene). The results showed that healthy people with the rare allele of the DARPP-32 polymorphism exhibited a better *Go* learning (choose A trials), whereas people with the rare allele of the C957T polymorphism exhibited a worse *NoGo* learning (avoid B trials). On the other hand, the Val^{158}Met polymorphism seemed to have an effect on people's ability to adapt their behaviour. A more recently study from the same group (Doll, Hutchison, & Frank, 2011), also confirmed that polymorphisms within the striatal dopaminergic genes may influence the *Go* and *NoGo* learning.

Another interested RL task was the gambling task developed by Daw and his colleagues (Daw *et al.*, 2006). In this study, the authors analysed the explore/exploit dilemma using the gambling task in an fMRI environment. During this task, in each trial, subjects should choose one slot-machine, out of 4 available slots, and then they received a feedback. The results indicated that both frontopolar cortex and intraparietal sulcus areas were active during the exploratory decisions, and that the striatum regions and ventromedial prefrontal cortex might be related with the value of the decisions of the exploitation.

More recently, the same group also adapted the gambling task, replacing the slot-machines by faces, to investigate the relationship between the RL and the hippocampal contextual memory (Wimmer, Daw, & Shohamy, 2012). Due to the generalization needed to make relational representations to solve this task, the modelling results suggested a possible functional connectivity between hippocampus and striatum structures.

Palminteri and his colleagues (Palminteri *et al.*, 2009) designed a new binary selection task, where subjects received a probabilistic feedback depending on their choice. Interestingly, in this fMRI study, they demonstrated that both intrinsic values of each choice (left or right) option were represented, respectively, in the contralateral area of the ventral prefrontal cortex. In addition, this group also applied this same task in adult patients with TS and verified that both simple motor tics and comorbid obsessive-compulsive disorder (OCD) were resultants of the dysfunction of the dopaminergic system (Worbe *et al.*, 2011).

Globally, all of these evidences have contributed to have a better understanding of the global RL process in humans. In addition, these tasks clearly demonstrated their applicability to study cognitive functions in the future.

3.2. Influence of age in reinforcement learning process

Several studies have also been developed to investigate how age can influence the RL process, on one hand, across neurodevelopment (Shephard, Jackson, & Groom, 2014; Baldwin *et al.*, 2012, Crone, Jennings, & Van der Molen, 2004), and on the other hand, in aging (for review: Eppinger, Hämmerer, & Li, 2011; Eppinger *et al.*, 2013).

A major study, which included 179 subjects: 44 children (9–11 years), 45 adolescents (13–14 years), 46 younger adults (20–30 years), and 44 older adults (65–75 years), aimed to monitoring the age-dependent differences in a probabilistic RL (Hämmerer *et al.*, 2011). The authors found that amplitude decrease of the feedback-related negativity across ages. Moreover, adolescents and younger adults versus children and older adults needed more trials to learn the conditions and learned less from positive than negatives feedback.

Latter, the same group (Eppinger *et al.*, 2013) described some differences between 13 younger ((mean age = 28.8) and 13 older (mean age = 70.0) adults using a RL selection task in a fMRI context. In this task, actions were reinforced by monetary rewards or losses, where the subjects needed to learn some of them and avoid others, respectively. The results indicated age-related decreased in learning for positive feedback (and a maintained learning from negative ones), which may seem dependent of a decline in phasic dopamine signalling in striatum.

A more recent study that included 14 children (mean age = 10.2) and 15 healthy adults (mean age = 25.5), using a task where visual stimuli leaded to the learning of a response through feedback, and then a re-learned due the un-expected reverse of those feedback. Children showed difficulties in situations where the acquired behaviours needed to be adapted to a new scenario, which the authors considered that a possible cause may be a highly demanded task for an immature executive systems, or a existent difference between the RL approach between children and adults (Shephard, Jackson, & Groom, 2014).

To finish, Jones and her colleagues (Jones *et al*., 2014) designed a robust study to better understand the patterns of behaviour in adolescents during social learning. This study included 120 subjects (8-25 years old), whose 68 completed the probabilistic social RL task in a fMRI environment. As already indicated, adolescents learned worsted from positives feedback when compared to both children and adults.

Aims

The RL cognitive tasks have shown to be important tools to study the RL process in humans. A well-characterized task is essential to better understand these aspects not only in healthy people, e.g. to study the influence of a specific genetic variations in behaviour (Frank & Hutchison, 2009; Frank & Fossella, 2010), but also in people with some neurologic and psychiatric disorders, who may have dopaminergic impairment and/or aberrant basal-ganglia circuitry (Maia & Frank, 2011).

Additionally, in the future, some behavioural tasks can be used as a clinical routine to better characterize the psychological profile of patients, offering a valid opportunity to better discriminate several endophenotypes of some disorders. In fact, this aspect is of great interest to the clinical practice, since a better understanding of some psychiatric disorders in the future, namely their endophenotypes, will allow refining both behavioural therapy and pharmacological treatment (Nemoda, Szekely, & Sasvari-Szekely, 2011).

In a previous study, Dias developed a probabilistic RL Go/NoGo task, which proved to be sensitive to both positive and negative reinforcers as well as to the motor biases in healthy humans (N = 24; 19-50 years old) (Dias, 2014). Therefore, the next logical step was to test whether this tool could be also used to study these aspects in several disorders, namely those that might have impairments in basal-ganglia circuitry. Furthermore, since there is a huge interest in applying it to the study of neurodevelopment disorders, such as TS, Attention deficit and hyperactivity disorder (ADHD), and OCD, which affect mostly people at younger ages (Nemoda, Szekely, & Sasvari-Szekely, 2011), it became evident the importance of testing such task in younger people.

In the light of these future interests, the main propose of this thesis is to **study the influence of age on the reinforcement learning process in healthy humans**, by applying a new version of the previously studied Go/NoGo cognitive task and a final questionnaire.

This new task version was developed from the former one with slight alterations, which intended to facilitate its correct understanding and execution when applied to both children and teenagers. Hereupon, it is expected to obtain behavioural data that clearly exhibits an active RL process throughout the task. Furthermore, since it is also expected that behaviour might be related to the prediction errors, the task design also took into account the possibility of a future application in studies of fMRI with both healthy people and patients.

In addition, the final questionnaire will also offer the opportunity of collecting more detailed demographic information and to have access to some data about the subjects' task execution. The latter will possibly allow having a better perception of the subjects' consciousness level of their own performance, as well as other aspects that might influence their performance.

To summarize, the specific aims of this work are the following:

1. To examine the influence of age in both RL process and motor biases in humans:
   a) Mathematical modulation of the behavioural data;
   b) Evaluation of possible differences within the learning of the task conditions, which are mainly related to the people's sensitivity to both positive and negative reinforcers.
2. To evaluate the people's consciousness level of their own task performance.

Finally, this study involves 3 main steps to accomplish all of the proposed aims: 1) the recruitment of healthy participants, whose age ranges from 6-80 years; 2) the application of the new probabilistic RL Go/NoGo task, followed by a final questionnaire to all participants; and 3) the statistical analysis of the obtained data.

Methodology

This chapter includes all methodologies and techniques that were used to accomplish this study, being divided into three sub-sections: 1) subjects recruitment; 2) materials used: Go/NoGo task and final questionnaire; and 3) statistical data analyses.

1. Subject recruitment

1.1. Subject sample

This study included 419 people, whose age ranges from 6-80 years. People were excluded if they had a lifetime history of any neuropsychiatric disorder (DMS-V), if they were currently taking or had taken in the last 4 months any psychiatric or neurologic medications, if they had any prior seizure, a history of head trauma with loss of consciousness, and birth before 36 weeks gestational age. All the considered people were Portuguese speakers.

1.2. Recruitment process

The recruitment process occurred during November of 2014 and February of 2015. Both children and teenagers, whose age ranged from 6-18 years old, were recruited mainly from three public schools: *Escola Básica 1º ciclo e Jardim de Infância do Miratejo* (EB1/JI), *Escola Básica 2, 3 de Corroios* (EB23C), and *Escola Secundária João de Barros* (ESJB). These educational institutions were included in the same cluster of schools, *Agrupamento de Escolas João de Barros* (Corroios, Portugal), which means that their students usually follow the same academic path.

Adults, whose age ranged from 18-80 years old, were recruited mainly from: 1) the academic environment, people whose age ranged from 18-30; and 2) a cultural and recreational local center, the *Centro Cultural e Recreativo do Alto do Moinho* (CCRAM) (Corroios, Portugal), people whose age ranged from 30-80.

The recruitment process included two main steps. First, the study was presented to potential participants, where a brief explanation of the task and questionnaire was made, including their durations and all exclusion criteria were carefully explained. Importantly due to the nature of the Go/NoGo task, this task was always presented to subjects as a “computer game”.

In addition, children from both EB1/JI and EB23C, whose age ranged from 6 to approximately 12 years old, received an additional verbal incentive to make this experience the more enjoyable possible: it was told to them that they were in a competition where the first three best results of each class would receive a “big” surprising prize, and the remaining participants would receive a “small” surprising prize.

Then, to people interested in joining the study, both information sheet and consent form were delivered (appendice A1 and A2 – page 123). Children and teenagers received two different information forms and consents forms (appendice A1 – page 123), to guarantee that both participant and his legal representative gave us permission to participate in the study. The adults only received one information form and one consent form (appendice A2 – page 129). All participants included in the study had delivered their consent forms before the time that they performed the cognitive task and answered the questionnaire.

2. Materials used: Go/NoGo task and final questionnaire

The Go/NoGo task and the final questionnaire were applied during a single session. In general, 3 to 12 people participating in each session.

Moreover, the enrolled participants received an identity number to be used in both task and questionnaire, in order to protect their anonymity and the confidentiality of data during its analysis, being each identity number coded as a unique numeric sequence of 8 digits.

2.1. Probabilistic RL Go/NoGo task

As mentioned previously, the RL paradigm used was adapted from a previously tested version of the RL Go/NoGo task (Dias, 2014). This Go/NoGo task was similar to a computer game, which main goal was to gather as many points as possible.

During the task, five different images were displayed on the computer screen, one at a time, and the participant needed to choose to press or not to press the space bar key of the computer keyboard (do a *Go* or a *NoGo* action, respectively), in order to win or to avoid losing points. At the end of each trial, a feedback message appeared on the screen. The possible feedback messages were: +1 (win one point), -1 (lose one point), and 0 (neither win nor lose points). To be better perceived, these messages appeared to be coloured in green, red and black, respectively.

All trials followed the same pattern, showed in the figure 7: 1) the blank screen (during 3000 ± 2000 ms); 2) the fixation point (during 1000 ms); 3) the stimulus was presented, which was one of the five possible images; 4) the loading screen (during 3000 ± 1000 ms); and 5) the feedback message, which could be either +1, -1 or 0 (during 1000 ms).

The difference between *Go* and *NoGo* trials relied on the participant's decision of pressing or withhold the space-bar key in order to perform a *Go* or a *NoGo* trial, respectively.

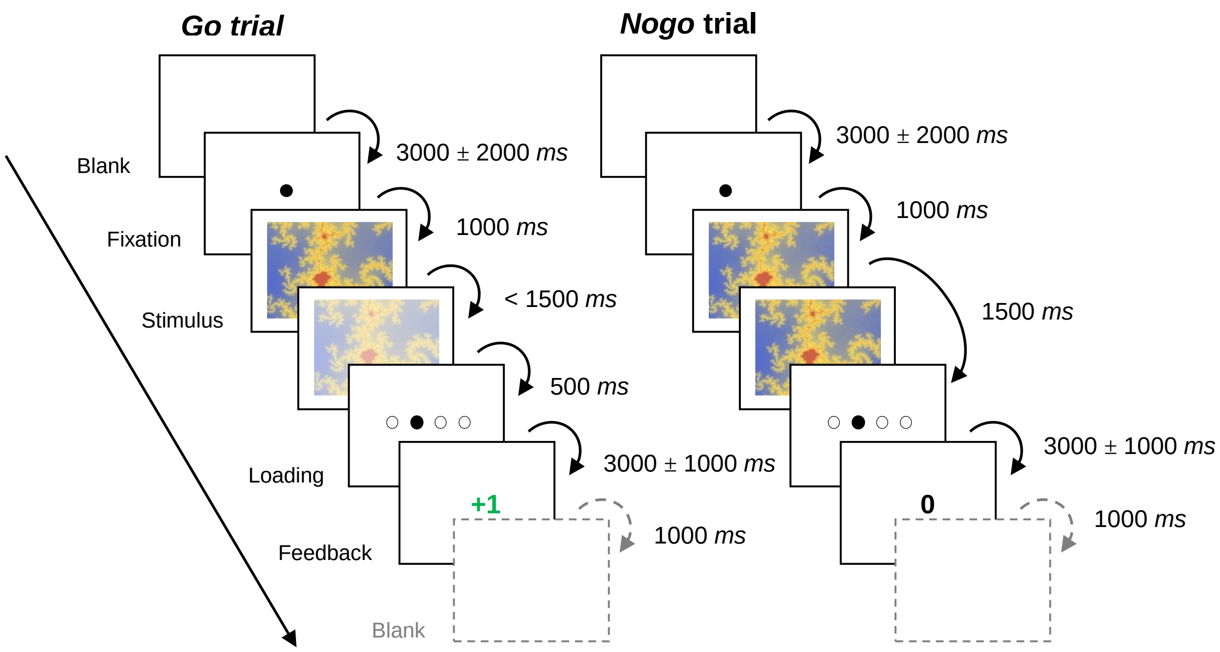


Figure 7. Schematics of both types of trials included on the task. Left side - *Go* trial; Right side - *NoGo* trial.

Moreover, each image had a displayed maximum-duration of 1500 ms, which was the available time for the participant decide his action. If the participant executed a *Go* trial, the image appeared translucent on the screen over an additional period of 500 ms (*Go* trial of figure 7), ensuring that the participant knows that his action was valid. Thus, the *Go* trial only occurred when the participants pressed the key while the image was on the screen, so every time that they pressed it outside this period their actions were not considered as a valid *Go* trial.

There was a total of five images (figure 8), one for each of the five conditions, which were randomly attributed to each participant: 1) *Go to win*, 2) *Go to avoid losing*, 3) *NoGo to win*, 4) *NoGo to avoid losing*, and 5) *Neutral*.

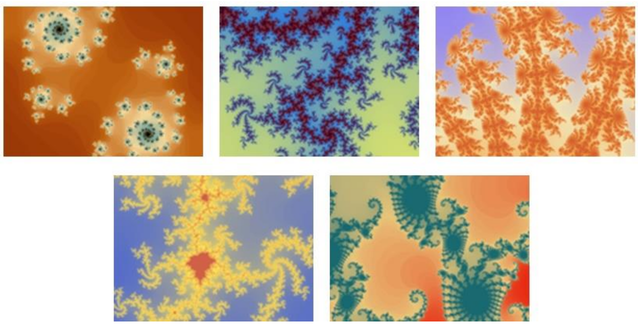


Figure 8. The five images used in the task.

The task was divided into three blocks, wherein each image/condition was randomly displayed 10 times per block, in a total of 30 times along the task, so the task had a total of 150 trials.

To make this task non deterministic, each condition (except the *Neutral* condition) had some probability of winning and losing points depending of which action was performed. The probabilities of the feedback message for each condition are presented in table 1. Therefore, to win points, the participants may learn to press and withhold the space-bar key in the *Go to win* and *NoGo to win* conditions, respectively; and to avoid losing points, to press and withhold it in the *Go to avoid losing* and *NoGo to avoid losing* conditions, respectively. In the *Neutral* condition, both actions gave the same feedback (0 points). In short, to obtain the maximum points possible, the participants need to learn to press the key in the *Go* conditions and withhold it in the *NoGo* conditions.

Furthermore, to ensure that the participants enrolled for this study really understood the Go/NoGo task before they begun to perform it (e.g. its aim, what key to press and when pressing it is considered valid), several strategies were applied. In fact, globally, the application of the Go/NoGo task included three phases: demonstration, training, and testing phases, being the last the Go/NoGo task itself.

Table 1. The probabilities of the feedback messages (-1, 0 and +1) of each action (*Go* and *NoGo*) per condition.

	Action					
	Go			***NoGo***		
Feedback / **Conditions**	**-1**	**0**	**+1**	**-1**	**0**	**+1**
Go to win	10 %	10 %	**80%**	0 %	**100 %**	0 %
Go to avoid losing	0 %	**100 %**	0 %	**80 %**	10 %	10 %
NoGo to win	0 %	**100 %**	0 %	10 %	10 %	**80%**
NoGo to avoid losing	**80 %**	10 %	10 %	0 %	**100 %**	0 %
Neutral	0 %	**100 %**	0 %	0 %	**100 %**	0 %

Therefore, before the participants begun to perform the Go/NoGo task they had to: 1) read (or listen the instructor to read) the writing instructions that appeared on the screen; 2) listened the additional verbal instructions given by the instructor; 3) paid attention to a demonstration phase; and 4) performed a training phase. Between each of these steps, the participants were encouraged to present their doubts, being the instructor responsible for moving forward in this process when he was fully satisfied with the participants' feedback. It is also worth noting that the presented phases did not have any other purposes besides helping the participants to understand the task. The next sub-sections describe each of them in more detail.

2.1.1. Instructions

The cognitive task always starts with some writing instructions (appendice A3 – page 133), which are divided into three parts. In situations that the youngest participants did not know how to read yet, the instructor read the instructions out loud. After that, the instructor explained again the main rules of the task, additional some complementary verbal instructions were given (appendice A4 – page 134), as well as all existing doubts among the participants were answered.

In addition, the instructor also informed the participants that all tasks were different from each other, explaining that even if equal images appear on the screen of two different games, the best action for each one of them will probably be different. This information was extremely necessary in order to avoid that the attention of the closest participants were lost in other tasks other than their own, which might deeply compromise the learning process.

Then, the instructor explained to participants that the following phase was just a demonstration of the task, while after they will also have the opportunity of trying the task by themselves during their own training phase.

The instructor also refereed that the images displayed during both of these phases (figure 9) were not the same used in the testing phase of the task, and that winning and losing points during these phases did not matter to their final performance, which only can be tested during the testing phase.

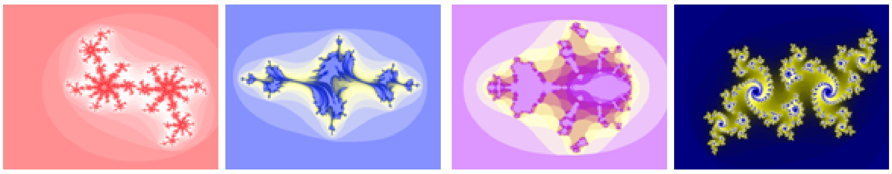


Figure 9. The four images used in both demonstration and training phases of the task.

2.1.2. Demonstration phase

After the instructor ensured that all participants understood the instructions of the Go/NoGo task, he performed the demonstration phase. This phase was equal for all participants and it included a total of 4 different trials, presented in table 2. In this phase, the instructor should performed a sequence of a *Go*, *NoGo*, *Go* and *NoGo* actions, in order to encompass all possible scenarios of winning and losing points during the task.

Furthermore, this phase was also essential to show the task dynamic (figure 7), and to highlight the difference between a *Go* and a *NoGo* trial: in a valid *Go* trial, the image appeared translucent for a brief instance after the instructor pressed the key, which needed to occur while the image was still displayed on the screen.

Table 2. Feedback messages (-1, 0, +1) of each trial in the demonstration phase, which only depends on the executed action (*Go* and *NoGo*).

	Action	
Trial	***Go***	***NoGo***
1st	+1	0
2nd	0	-1
3rd	-1	0
4th	0	+1

2.1.1. Training phase

The training phase was essential so that participants had the chance of “trying” the task before feeling the pressure of the possibility of losing and winning points.

Similar to the demonstrating phase, this phase allowed a better understanding of the task dynamic, namely the duration of each presented element of the trial sequence. It also offered an opportunity for participants to perceive how much time they had to look to the image, and decide what action they wanted to execute and to press the key in the cases that they chose performing a *Go* trial.

In contrast to the demonstration phase, the training phase was not equal to all participants, since the images (figure 9) were randomly attributed to the 4 considered conditions: *Go to win, NoGo to win, Go to avoid losing*, and *NoGo to avoid losing*. Moreover, each image was randomly displayed three times so that this phase had a total of 12 trials. Similar to the demonstration phase, the feedback of each condition only depended on the executed action, since there was no probabilities involved, which are presented in following table (table 3).

Table 3. Feedback messages (-1, 0, +1) of each considered condition in the training phase, which only depends on the executed action (*Go* and *NoGo*).

	Action	
Trial	**Go**	*NoGo*
Go to win	+1	0
NoGo to win	0	+1
Go to avoid losing	0	-1
NoGo to avoid losing	-1	0

2.2. Final questionnaire

After performing the task, all participants should answer to the final questionnaire (appendice A5 – page 135). This questionnaire was developed using the *Google Forms* platform, so it became available on-line and answering it presupposed an internet connection. It was fully filled by the oldest participants, but the youngest participants received a closely monitoring by the instructor, which also included filling the questionnaires with the answers given by the participants younger than 10 years old.

The questionnaire was organized in two parts: the first included some demographic questions (1st to 17th question) and, the second, some questions about the task performance (18th to 40th question).

2.2.1. Demographic questions

In summary, the questions of the first part of the questionnaire allowed to better characterize the participants' sample (e.g. participant's birthday date, his gender, and his professional situation). It is worth noting that this part of the questionnaire was constructed in a way that the subjects' answers influence the amount of questions to be answered along the questionnaire, having the questionnaire a total of 17 questions in this part.

In fact, subjects whose school degree was between 1st and 9th year of the basic education, whose age-range is usually between 6 and 15 years old, only had to answer to 8 questions. This number increased if the subjects' school degree was between 10th and 12th year, being their age-range usually between 16 and 18 years old: if they were doing a CEF course (or *Curso de Educação e Formação)*, they needed to answer to 9 questions; and if they were doing a professional or a general course (or *Curso Científico Humanístico*), they needed to answer to 10 questions. Participants that were under- or graduated students also needed to answer to 10 questions. Finally, the participants who were no longer studying, usually the oldest participants, only needed to answer to 9 questions.

2.2.2. Questions about the task performance

The second part of the questionnaire addressed questions about the execution of the task by the participants: 1) theirs self-evaluations of their attention during the task, 2) their classifications about the beauty of each image, 3) the action that they considered the best to perform for each image, and 4) which points were frequently showed after they performed a *Go* or a *NoGo* action in each image.

After the demographic questionnaire's part, and before the participants begun to answer to the enumerated questions, the questionnaire included a transitory information sentence displayed on the screen: "*As próximas questões são acerca* *do jogo que acabaste de jogar* / The following questions are about the game that you just finished to perform". This part included a total of 23 questions.

- **Attention during the task**

Both 18th and 19th questions of the questionnaire were about the subjects' attention during the task performance.

The 18th question: "*Achas que estiveste sempre atento durante o jogo?* / Do you think that you were paying attention during the entire game?", had two possible answers:

- "*Sim, estive sempre atento durante todo o jogo* / Yes, I was paying attention during the entire game";
- "*Não, nem sempre consegui estar atento durante todo o jogo* / No, I could not always pay attention during the game".

In addition, if the participant answered "no" to this question, he automatically needed to answer to the 19th question, which was "*Em que parte(s) do jogo não estiveste atento?* / In which part(s) of the game did you not paid attention?". This question given the opportunity of chosen in which parts of the task (1st, 2nd, 3rd and/or during the entire task) he felt that he was not paying attention to his performance.

- **Amount of different task's images**

The 20th question of the questionnaire: "*Quantas imagens diferentes achas que o jogo tinha?* / How many different images do you think that the game had?", directly challenged the participant to digit the number relative to his answer.

- **Images' beauty classification**

This questionnaire's section also started with a transitory information sentence saying: "*As próximas questões são apenas referentes à beleza de cada imagem. Não tem nada a ver com o que fizeste no jogo.* / The following questions are only about the beauty of each image. It is not related with what you did during the game".

Therefore, the next 5 questions (21st, 22nd, 23rd, 24th, and 25th) included one of the task's images followed by the rating scale: "Muito feia / Very ugly", "Mais ao menos feia / More or less ugly", "Nem feia nem bonita / Neither ugly nor beautiful", "Mais ao menos bonita / More or less beautiful", and "Muito bonita / Very beautiful". The participants could only choose one answer for each image.

- **Best action and points received for each image**

This section begun with the following message: "As próximas questões são referentes ao que fizeste no jogo. É normal que demores algum tempo para te recordares daquilo que fizeste. / The following questions are about of what you did in the game. It is normal that you take some time to remember of what you did.".

Similarly to the previously section, in this section, for each displayed image it was also presented three questions. The first: "Nesta imagem, o que era melhor fazeres? / In this image, what was the best action to perform?", had 4 possible options:

- "*Carregar na barra de espaços* / Press the space-bar key";
- "*Não carregar na barra de espaços* / Do not press the space-bar key";
- "*Tanto fazia carregar ou não na barra de espaços* / It did not matter pressing or not the space-bar key";
- "*Não sei / Não me lembro* / I don't know / I don't remember".

Since each image allowed the execution of both *Go* and *NoGo* actions, the 2nd and 3rd questions distinguished the points received concerning those actions, respectively:

- "*Se CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?* / If you PRESS the space-bar key when this image appeared, which points were usually displayed on the screen?";

- "*Se NÃO CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?* / If you DID NOT PRESS the space-bar key when this image appeared, which points were usually displayed on the screen?".

To both questions the possible answers were "-1", "0", "+1", "*Não sei / Não me lembro* / I don't know / I don't remember".

3. Statistical analyses

The statistical methodologies used during this study are included in this section, which is organized in 3 main sub-sections: 1) sample characterization; 2) Go/NoGo task data; and 3) questionnaire data. Furthermore, these analyses were performed by using the software *R* version 3.0.2, through the *RStudio* version 0.98.501 (© 2009-2013 RStudio, Inc), and the *Microsoft Office Excel 2007* software version 12.0.

3.1. Sample characterization

As mentioned previously, the information about the profiles of the participants were collected by the first part of the final questionnaire. That being said, this section included a global sample characterization, which included all participants. Then, due to the participants' great range of ages (from 6 to 80 years old), they were divided into different clusters, allowing to verify possible effects due to their age.

Therefore, to better understand the RL process across age, the subjects' sample were divided into 18 smaller groups: [6.0-7.0), [7.0-7.5), [7.5-8.5), [8.5-9.5), [9.5-10.5), [10.5-11.5), [11.5-12.5), [12.5-13.5), [13.5-14.5), [14.5-15.5), [15.5-16.5), [16.5-17.5), [17.5-18.5), [18.5-20.5), [20.5-24.0), [24.0-30.0), [30.0-50.0), [50.0-80.0)[1] years old. To simplify, these groups will be identified, respectively, as: 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19-20, 21-23, 24-29, 30-49, 50-80 years old.

3.2. Go/NoGo task data

The statistical analysis of the data collected by the subjects' executions of the Go/NoGo task was divided into two main analyses. The first part analysed the participants' performance through the entire task, by making use of the learning curves and the proportion of correct action computed for each block. The second

[1] Note that the type of the presented intervals, [x,y), represents an age-range between x and y, which includes the exact value of x and excludes the exact value of y from the age-range.

part intended developing a statistical model, which used the proportion of correct action of the last 20 trials relative to each task condition. This aimed to mathematically explain the behavioural performance across age.

3.2.1. Learning curves

In order to obtain the learning curves across trials, the individual probability of the *Go* actions for each condition was averaged across subjects. Furthermore, the learning curves for each age cluster were also computed. These were smoothed by a central moving average filter with a length of 5 and they were depicted across trial.

Globally, these analyses allowed a better understanding of the RL process, namely if the learning of each condition depended on the subjects' ages.

3.2.2. Participants' scores

All subjects' scores were calculated by adding the gained points and subtracting the loosed points during the entire task performance (trial number = 150), and during both 1st and 2nd-half of the task. Then, the averages of these calculated scores were computed by subjects' groups to verify if there was a tendency between the learning process and the obtained scores across age.

3.2.3. Proportions of correct actions

The task performance was measured through the proportion of correct action (*Proportion (CA)*) in each block. In addition, the considered correct actions were: the *Go* action for both *Go to win* and *Go to avoid losing* conditions; and the *NoGo* action for *NoGo to win*, *NoGo to avoid losing*, and *Neutral* conditions. Although both actions in the *Neutral* condition neither gave nor took points, the correct action to perform was also considered the *NoGo* action, since it is the less effortful action. The correct actions for each condition are presented in table 4.

Table 4. The correct actions considered for each condition.

Conditions	Correct action
Go to win	*Go*
Go to avoid losing	*Go*
NoGo to win	*NoGo*
NoGo to avoid losing	*NoGo*
Neutral	*NoGo*

The averages of the correct actions for each condition were computed across subjects of each group in order to verify possible differences among the conditions learning and the subjects' ages. In fact, learning a condition presupposed an improving performance during the task, which means a gradual increasing of the computed values across the different blocks of the task, and that these values (or at least the one corresponding to the last block) were higher than 0.5, which was the threshold value of performing the task by chance.

3.2.4. Behavioural model

The following analyses aimed to develop a mathematical regression (or model), that might explain the relationship between the task performances and the age, wherein the aforementioned individual proportions of correct actions for each subject were used as the dependent variable of the regression.

Moreover, this model was fitted to the behavioural data, taking into account the following explanatory variables of each condition: its correct action (action), valence, and neutral condition.

The values for these variables were binary (0 or 1), leading to that each condition had its specific code given by its variables' sequence, which are presented in table 5, being the *NoGo to avoid losing* condition the reference category.

Table 5. Binary code used to distinguished action, valence, and neutral as independent variables within each condition.

Conditions	Action	Valence	Neutral
Go to win	1	1	0
Go to avoid losing	1	0	0
NoGo to win	0	1	0
NoGo to avoid losing	0	0	0
Neutral	0	0	1

Furthermore, the tested regressions could be divided into two groups, having the first used the proportion of correct action (p), and the second one, the logarithm of the odds ratio ($\log\left(\frac{p}{1-p}\right)$ or $\mathrm{logit}(p)$) as the dependent variables. In addition, both used the subjects' ages (age) as an independent variable. These general regressions are respectively presented by the following equations:

$$\begin{aligned} p \sim\ & \beta_0 + \beta_1 valence + \beta_2 neutral + \beta_3 action + \beta_4 valence \text{ x } action \\ & + \beta_5 age + \beta_6 valence \text{ x } age + \beta_7 neutral \text{ x } age + \beta_8 action \text{ x } age \\ & + \beta_9 valence \text{ x } action \text{ x } age, \end{aligned} \tag{4}$$

$$\begin{aligned} \log\left(\frac{p}{1-p}\right) \sim\ & \beta_0 + \beta_1 valence + \beta_2 neutral + \beta_3 action \\ & + \beta_4 valence \text{ x } action + \beta_5 age + \beta_6 valence \text{ x } age \\ & + \beta_7 neutral \text{ x } age + \beta_8 action \text{ x } age \\ & + \beta_9 valence \text{ x } action \text{ x } age. \end{aligned} \tag{5}$$

In addition, the presented models were compared with the ones that did not include the age as an explanatory variable, in order to verify that the age was indeed an important variable to explain the obtained behavioural data.

Finally, several transformations were applied to the age variable: age, $\log(age)$, $\log(age) + age$, $\sqrt{age}$, age^2, and $age^2 + age$.

3.2.4.1. Model comparison

After having developed all the aforementioned models it became necessary to compare them. To do so, different methods were applied.

The log-likelihood function, or simply log-likelihood, can be used to compare models with the same number of parameters, where the best model is simply the one with highest log-likelihood value, since in that case no correction for different model complexities is necessary. On the other hand, both Bayesian information criterion (BIC) and Akaike information criterion (AIC) criteria can be used to compare nested models, since both criteria include correction terms to deal with the different number of parameters of the models. Both criteria are approximations of the model evidence and they should be used together, since the BIC over penalizes model complexity, and the AIC under penalizes it (Penny, 2012). Therefore, the AIC tends to select more complex models and the BIC the less complex ones. Ideally, both criteria should select the same model, being the best model the one that shows the lowest values for both criteria.

That being said, the first model comparison aimed to verify if the models that considered the subjects' age as an independent variable were better than those that did not include it, by using both AIC and BIC criteria.

After this step, and since the age was confirmed to be an important explanatory variable, it became essential to compare models with different variable transformations, using the log-likelihood function. In addition, other measures, as the squared-R (R^2), the adjusted-squared-R^2 (Adj-R^2), and the p-values were also used to confirm the selection of the best model.

After selecting the best transformation for both dependent (p) and independent (age) variables, it was the time to verify the importance of each of the others explanatory variables included in the selected model. With this purpose, two manual stepwise regressions were performed, considering either BIC or AIC criteria.

3.2.4.2. Contrasts

The idea of using contrasts came with the need of taking more precise conclusions about the behavioural data. This analysis considered the beta parameters of the developed linear model, being each condition defined by a specific cluster of beta parameters, which were called as the contrasts weights.

In this case, contrasts aimed to analyse the differences between the learning rates of the following presented aspects:

- Congruent versus Incongruent conditions, which depends on both conditions' action and valence. The congruent conditions are the *Go to win* (action = 1, valence = 1), and the *NoGo to avoid losing* (action = 0, valence = 0), and the incongruent ones are the *Go to avoid losing* (action = 1, valence = 0), and the *NoGo to win* (action = 0, valence = 1);
- *Go* versus *NoGo* conditions, which only depends on the action of the conditions;
- Win versus Avoid losing conditions, which only depends on the valence of each condition;

It is worth noting that these contrasts might use either slopes, intercepts or asymptotic values of the model regression for each condition.

3.3. Questionnaire data

As mentioned previously, the data of the first part of the questionnaire were used to characterize the participants' sample, and its second part provided data concerning the participants' task performances.

3.3.1. Attention during the task performance

The attention during the task performance was evaluated through two questions. The first question (18th question) implied a "yes/no" answer, depending on whether the participant considered that he paid or not attention during his task performance. The answers to this question were analysed using the absolute frequencies of each option for each subjects' group.

In addition, if the subject answer was "No", he automatically needed to answer to the next question (19^{th} question) about in which part(s) of the game he did not paid attention to his performance, being the subject free to selected more than one of the following 4 options: 1^{st} block, 2^{nd} block, 3^{rd} block, and during the entire task. In this case, the answers to this question were analysed by the relative frequencies of each option for each subjects' group.

Both of these analyses were conducted to verify if there were discrepancies within the subjects' groups or a major tendency shared within all participants to feel more difficulties during the execution of a specific part of the task.

3.3.2. Amount of different task's images

Since this question allowed any numeric answer, the absolute frequencies of all answers offered the possibility to verify the subjects' perception about the amount of different images included in the task. It may be need to cluster some answers, depending on the amount of different given answers, in order to facilitate this analysis.

3.3.3. Images' beauty classification

First of all, and to facilitate the following analyses, the qualitative scale used to rate the images' beauty, which consisted into 5 categories: very ugly, more or less ugly, neither ugly nor beautiful, more or less beautiful, and very beautiful; was converted into a quantitative scale, which ranged from 1 to 5 points, respectively.

Globally, the images' rating offered an important opportunity to verify if there was an equilibrate preference of all images in all age-groups to ensure the subjects' image preference did not bias the results.

Furthermore, this scale also aimed to understand if the condition behind each image affected the subjects' perception about its beauty. It is worth noting that the different rates of the scale could have a different intrinsic value for each subject. Therefore, this data was normalized within subjects using a Z-distribution, since this analysis also aimed to capture possible differences between images' conditions and their beauty. After this step, the averages were computed used the z-scores of each condition, which offered a more reliable measure of the subjects' preferences.

To verify if possible differences between subjects' preferences were statistically different (p-value < 0.05), the analysis of variance (ANOVA) was made followed by the Tukey post-hoc test.

3.3.4. Best action and received points for each condition

Similarly to the analyses of the images' beauty classification, both best action and points received for each image were interpreted according with the images' conditions. These two aspects were independently analysed.

Since the first presented aspect addressed to investigate the subjects' knowing/perception of which action was the best to execute given a specific image/condition, the questions had 4 possible answering options: "press the space-bar key" (*Go* action), "do not press the space-bar key" (*NoGo* action), "it does not matter pressing or not the space-bar key", and "I don't know / I don't remember"; their absolute frequencies were computed.

Both questions concerning the points received after executed a *Go* and a *NoGo* action were analysed in a similar way. In both questions, for each image/condition, the 4 possible answering options were: "-1", "0", "+1", and "I don't know / I don't remember". Therefore, the absolute frequencies of the subjects' answers were computed taking into account each condition.

Results

The main results of this study are included in this chapter, which is divided into: 1) sample characterization, 2) learning process, and 3) questionnaire's results. Additional results are also included in the supplementary results section (page 109) to complement the ones presented next.

1. Sample characterization

This study included 419 people, whose ages ranged from 6-80 years old: 17.36 ± 10.78 years[2]. Furthermore, 51.6% of the participants were males and 48.4% were females.

Figure 10 shows the participants distribution according to their groups, where each bar represents the absolute frequency of the subjects within each age-range, wherein the white and the grey sections represent the absolute number of the females and males participants, respectively.

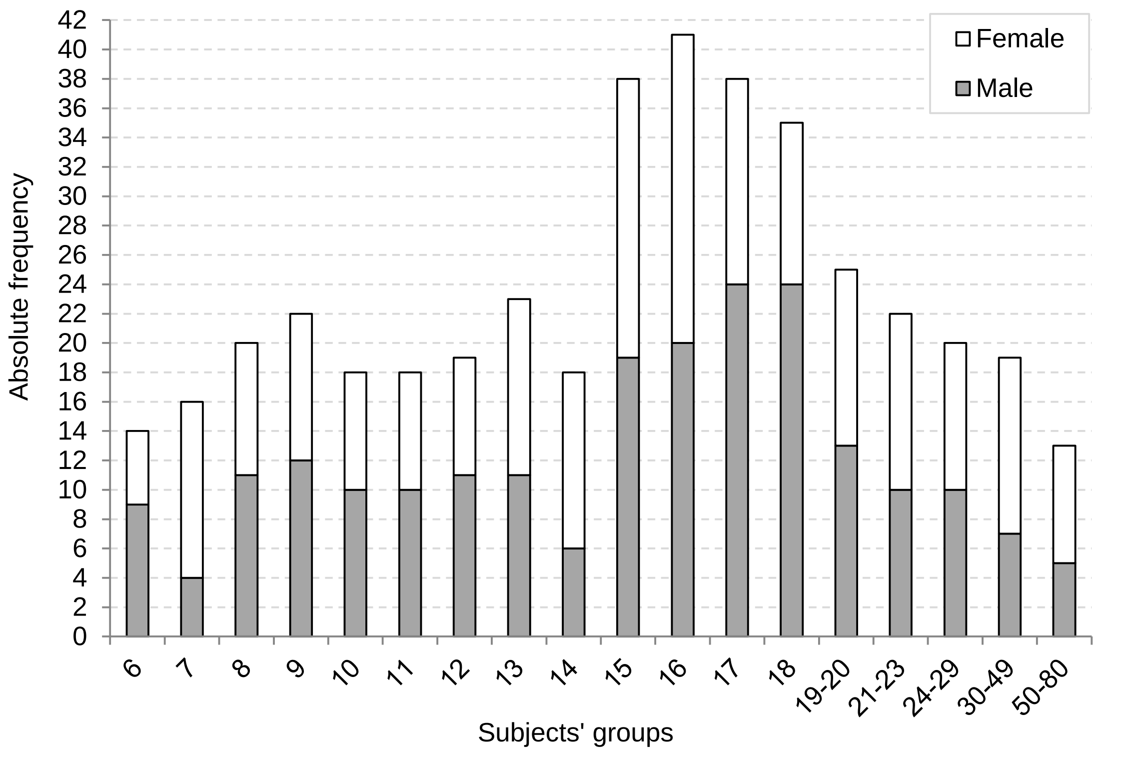


Figure 10. Participants' distribution according to their age-range, including its division in both female (orange/top section) and male (blue/bottom section) participants.

[2] Note that the "$x \pm y$" expression represents a numeric interval, being x its mean and y its standard deviation (SD). The interval may also be represented as $[x - y; x + y]$.

The following table (table 6) presents the averages of the considered age-ranges and the respective standard deviation (SD), both male and female frequencies and the absolute number of the participants of each subjects' group. On average, each subjects' group includes 20 ± 8.7 subjects.

Table 6. Characterization of each subjects' group: age-range mean, standard deviation (SD), both male and female frequencies (%), and the absolute number of participants.

Subjects' group	Subjects' age-range[1]	Age-range (mean)	Age-range (SD)	Male (%)	Female (%)	Subjects (N)
6	[6.0-7.0)	6.67	0.24	64.3	35.7	14
7	[7.0 - 7.5)	7.04	0.14	25.0	75.0	16
8	[7.5 - 8.5)	8.07	0.30	55.0	45.0	20
9	[8.5 - 9.5)	9.07	0.27	54.6	45.5	22
10	[9.5 - 10.5)	9.86	0.22	55.5	44.4	18
11	[10.5 - 11.5)	11.01	0.28	55.6	44.4	18
12	[11.5 - 12.5)	11.93	0.29	57.9	42.1	19
13	[12.5 - 13.5)	12.99	0.29	47.8	52.2	23
14	[13.5 - 14.5)	14.06	0.27	33.3	66.7	18
15	[14.5 - 15.5)	15.06	0.30	50.0	50.0	38
16	[15.5 - 16.5)	16.02	0.30	48.8	51.2	41
17	[16.5 - 17.5)	17.01	0.28	63.2	36.8	38
18	[17.5 - 18.5)	18.01	0.27	68.6	31.4	35
19-20	[18.5 - 20.5)	19.16	0.80	52.0	48.0	25
21-23	[20.5 - 24.0)	22.47	1.22	45.5	54.6	22
24-29	[24.0 - 30.0)	25.24	1.69	50.0	50.0	20
30-49	[30.0 - 50.0)	40.27	6.65	36.8	63.2	19
50-80	[50.0 - 80.0)	60.86	10.96	38.5	61.5	13
Total		17.36	10.78	51.6	48.4	419

2. Learning process

The RL process was evaluated taking into account the learning curves across all trials, the participants' scores, and the proportion of correct actions of each condition for each block. Finally, using the proportion of correct actions of the last 20 trials of the task of each condition, we developed a mathematical model and performed some contrasts.

2.1. Learning curves

In general, the learning curves, which are given by the *Go* probability for each trial number, offered an easier opportunity to visualize the learning evolution of each task condition.

Globally, the learning curves of the sample are presented in figure 11, where the final reached probabilities of the *Go to win*, *Go to avoid losing*, *Neutral*, *NoGo to win*, and *NoGo to avoid losing* conditions are, approximately, 0.80, 0.70, 0.48, 0.42, and 0.30, respectively.

In addition, the learning curves of each subjects' group are presented in both figures 12 and 13, where each graph represents data from one group.

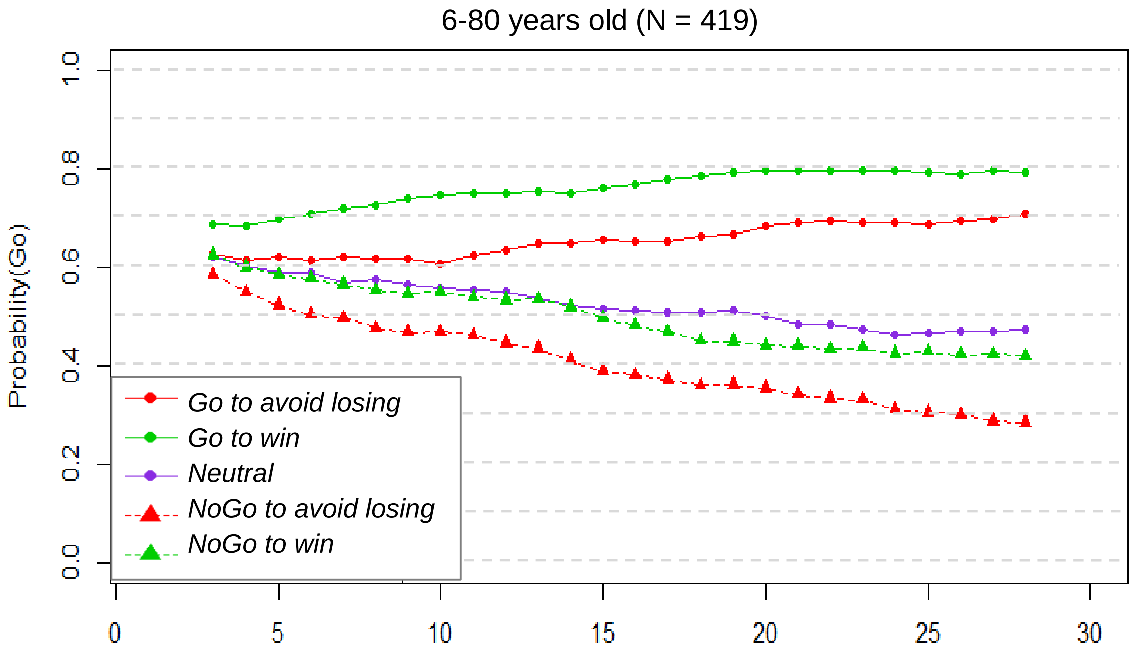
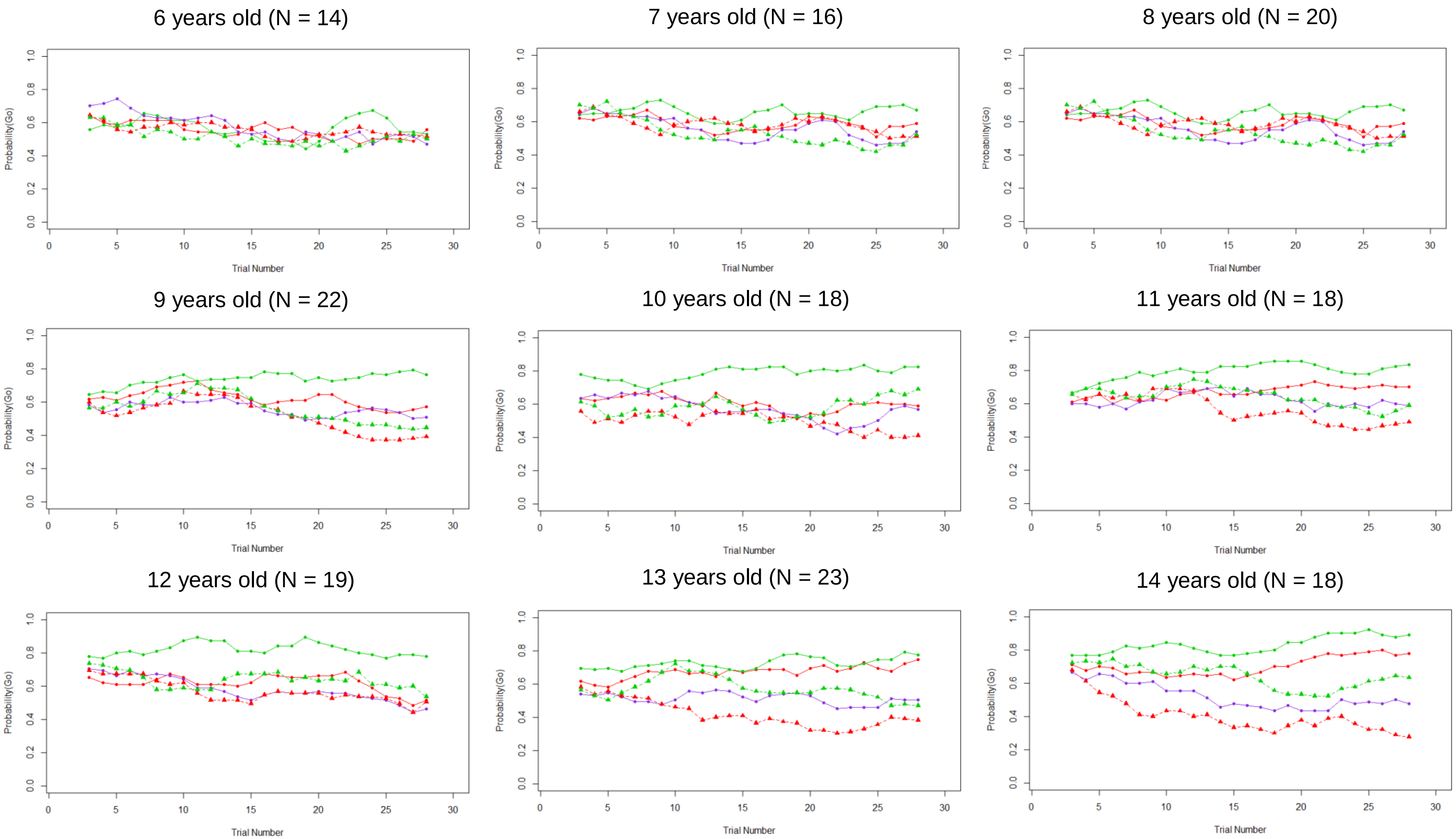


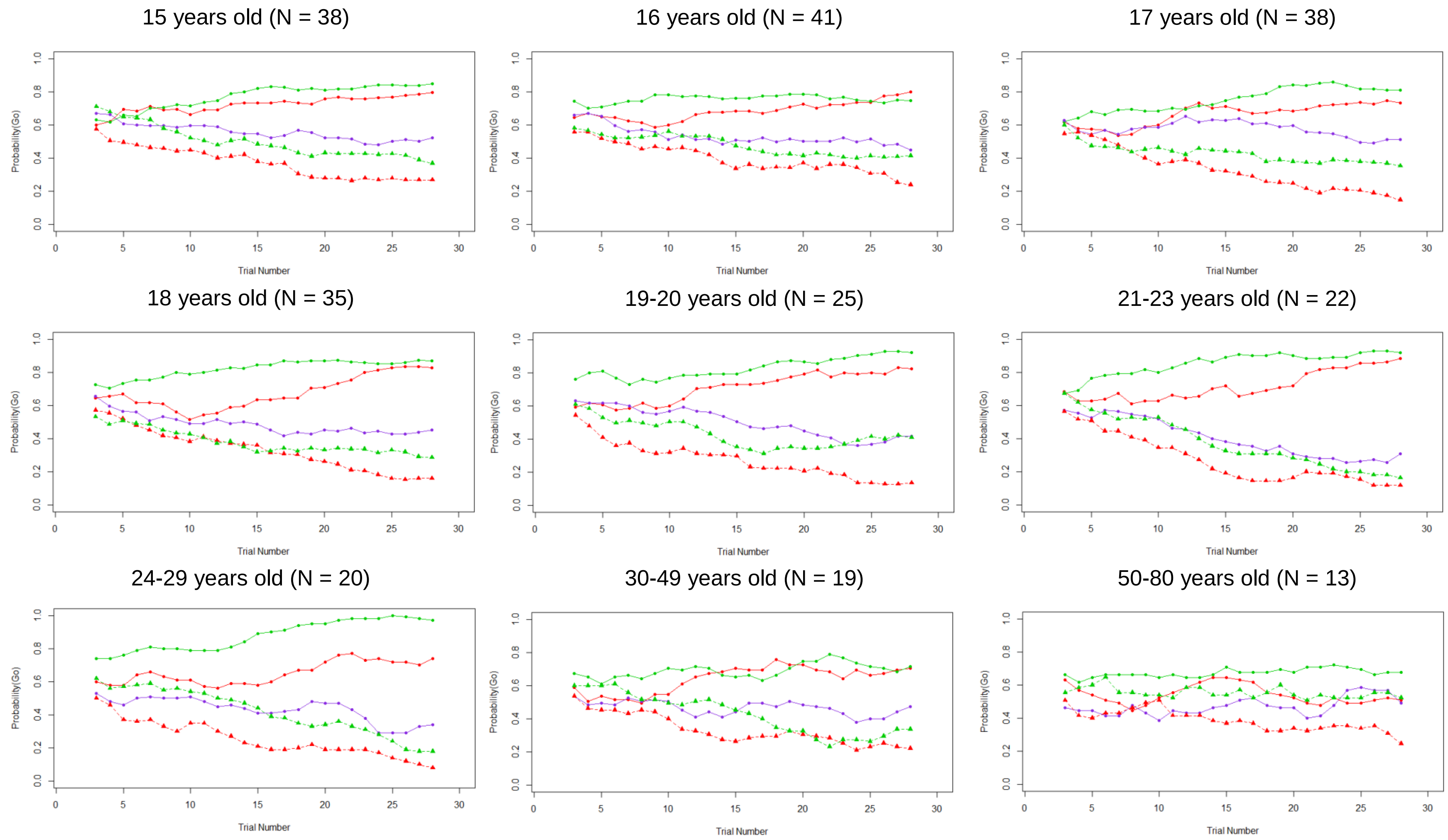
Figure 11. The Go/NoGo task's learning curves of all participants. Legend is presented on the right-bottom side of the graph.



Legend:

Go to avoid losing
Go to win
Neutral
NoGo to avoid losing
NoGo to win

Figure 12. The Go/NoGo task's learning curves. Each graph includes the learning curves of each subjects' group, being both group identity and its number of included subjects (N) are presented in the titles of the graphs. From left to right and top to bottom, each graph represents, respectively, the learning curves of 6, 7, 8, 9, 10, 11, 12, 13 and 14 years old subjects' groups; followed by the respective legend.



Legend:

Go to avoid losing *NoGo to avoid losing*
Go to win *NoGo to win*
Neutral

Figure 13. The Go/NoGo task's learning curves. Each graph includes the learning curves of each subjects' group, being both group identity and its number of included subjects (N) are presented in the titles of the graphs. From left to right and top to bottom, each graph represents, respectively, the learning curves of 15, 16, 17, 18, 19-20, 21-23, 24-29, 30-49 and 50-80 years old; followed by the respective legend.

As it can be observed, all of these graphs show a clear temporal relationship not only between the *Go* probabilities of each condition and the trial number, but also with the subjects' ages.

In general, the *Go* probabilities increase across trials for both *Go* conditions (*Go to win* and *Go to avoid losing*), reaching values above 0.8 for some subjects' groups; decrease for both *NoGo* conditions (*NoGo to win* and *NoGo to avoid losing*), reaching values bellow 0.2; and they appeared to be fairly constant for the *Neutral* condition, around 0.4-0.5. In fact, the 21-23 years-old subjects' group presents the highest and the smallest values of the *Go* probabilities for both *Go* and *NoGo* conditions, respectively.

Furthermore, the youngest subjects, from both 6 and 7 years-old subjects' groups, showed a task performance approximately equal for ever condition, being the *Go* probabilities a result of solving the task by chance. Then, each condition differentiates from the others, including from the *Neutral*, at different ages. The first conditions to stand-out were the congruent ones: first, the *Go to win* at the 9 years-old subjects' group, and then the *NoGo to avoid losing* condition showed a tenuous distinction from the others between 9 and 11 years-old subjects' groups, that becoming solid in the 13 years-old subjects' group. Both incongruent conditions, *Go to avoid losing* and *NoGo to win* were clearly separated at 13 and 15 years-old subjects' groups, respectively.

2.2. Participants' scores

Another way to measure the subjects' performances was by the analysis of the obtained scores after the task execution. Once subjects had the aim of trying to gain the maximum points possible, it is expected that the total subjects' scores might be directly proportional to their performances.

Figure 14 represents not only the participants' scores of the entire task but also the scores of both 1st and 2nd-half of the task (75 trials each). As it can be observed, the total subjects' scores of the entire task (blue bars) show an inverted-U shape, being the maximum reached around the 21-23 years-old subjects' group.

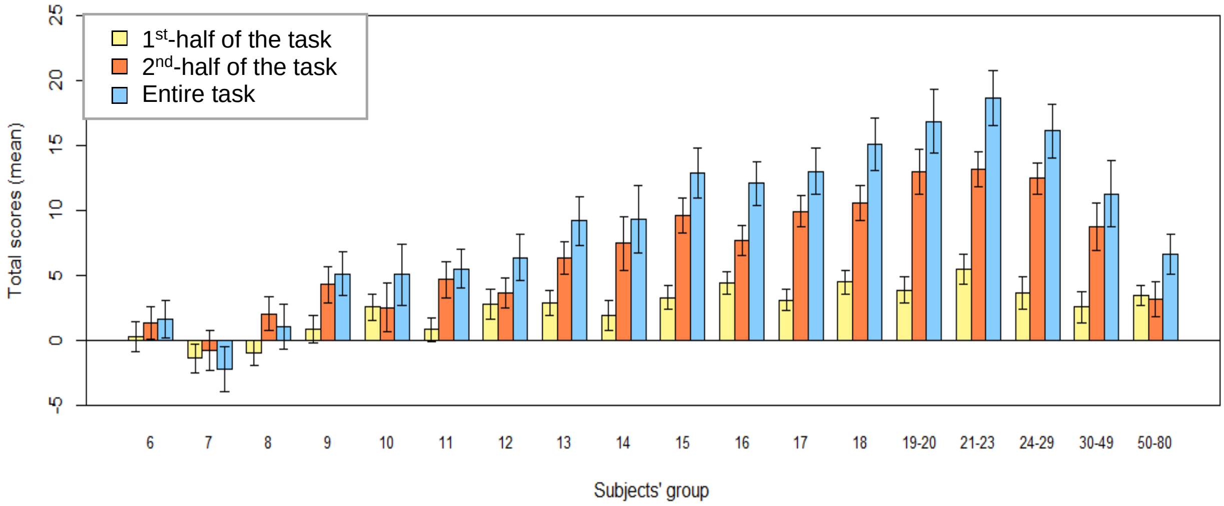


Figure 14. Average of total scores of both 1st and 2nd-half of the task, as well as of the entire task, for each subjects' group. Legend on the right-top side of the graph.

Moreover, the total scores of the 2nd-half of the task are, on average, greater than the ones of the first-half (except to the 6, 7 and 50-80 groups). It is important to noting that the total score of the 7-years-old subjects' group stand out by being always negatives on average.

2.3. Modelling the behavioural data

This section includes all the results considered important to achieve the model which best explains the subjects' performance across age. It is worth noting that this model was develop aiming to mathematically explain the relationship between the proportions of the correct actions of the last 20 trials of each condition and the subjects' ages.

2.3.1. Excluded participants

Since the following analyses aimed to model the learning behaviour of the participants, the *Go* probabilities of all conditions were analysed to verify the possible existence of participants that either always pressed or never pressed the space-bar key during, approximately, the entire task. The amount of this kind of participants within the collected sample might provoke a bias into the results, where their misfit behaviour showed a learning of 0% in some conditions and a "perfect learning" (100%) in others.

Therefore, after this analysis, only 6 participants (from 419) showed this behaviour, being the outliers. This way, they were excluded from the following analyses. On the other hand, and since these subjects came from different groups, their exclusion slightly affected the sample distribution. Table S1 (page 113) includes the new characterization of subjects' groups used in the next analyses, being the affected groups highlighted in bold.

2.3.2. Proportions of correct actions

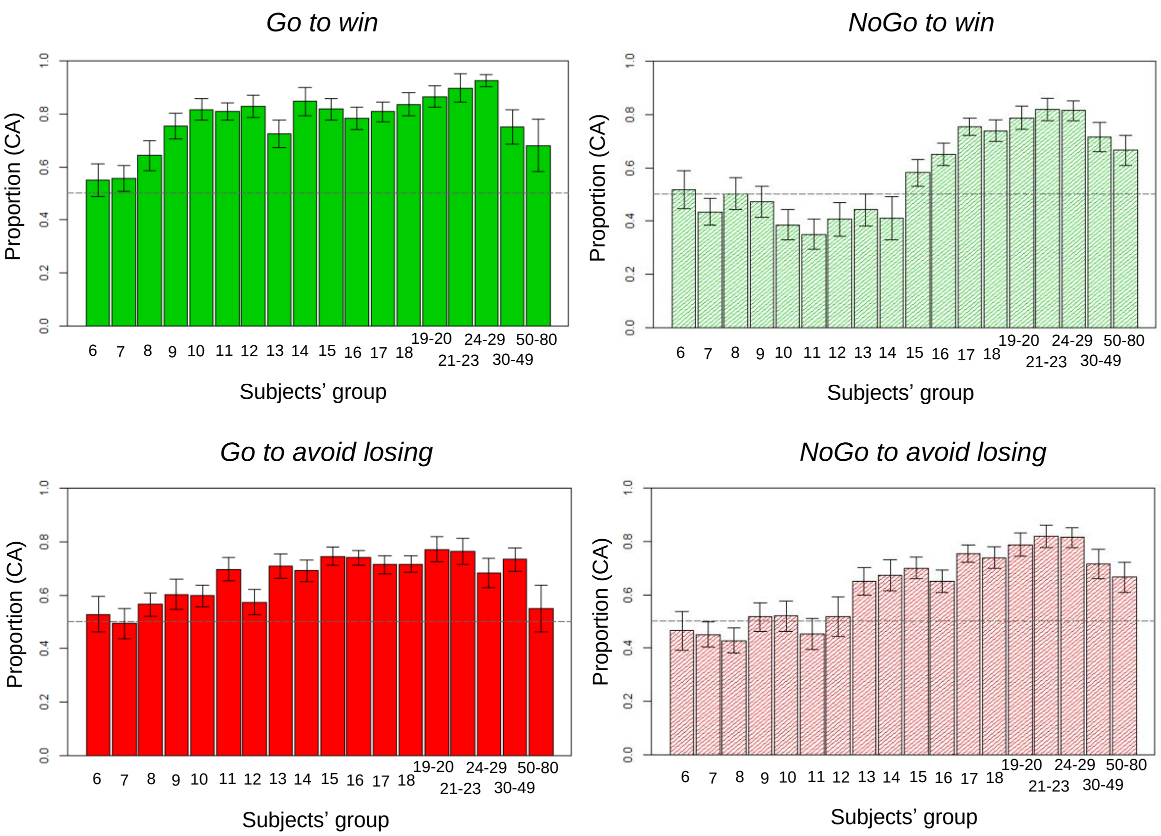
As mentioned previously, the overall performance of each subject was analysed by computing the proportions of the correct actions along the task for each task condition. Similarly, the average of the proportion per subjects' group for each condition offered a reliable measure to analyse the overall performance by age.

Since the model used the proportions of the correct action of each condition considering the last 20 trials of the task, these proportions were calculated and then clustered by subjects' group, in order to compute the respective averages considering each condition.

In addition, only the obtained data of subjects whose ages were less than 30 years old were used in this model. This cut-off relied on the fact that the oldest subjects (≥ 30 years old) were far less when compared to the youngest: it is worth noting that the ratio was 15.91 subjects/years ($382\, subjects/24\, years\ old$) for the youngest subjects' groups, and only 0.62 subjects/years for the oldest groups ($31\, subjects/50\, years\ old$).

The following barplots (figure 15) present the overall relationship between learning and age for each condition, considering the proportions' averages for the last 20 trials, having each graph information about one of the five conditions.

In this scenario, the learning of each condition depends on whether the computed proportions reach values clearly above 0.5 (horizontal dashed line), which implies that both averages and error-bars are above 0.5.



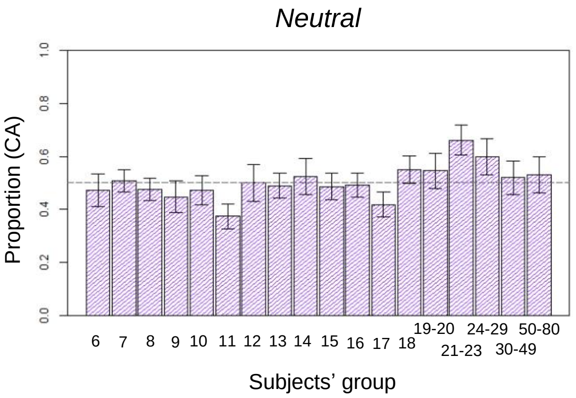
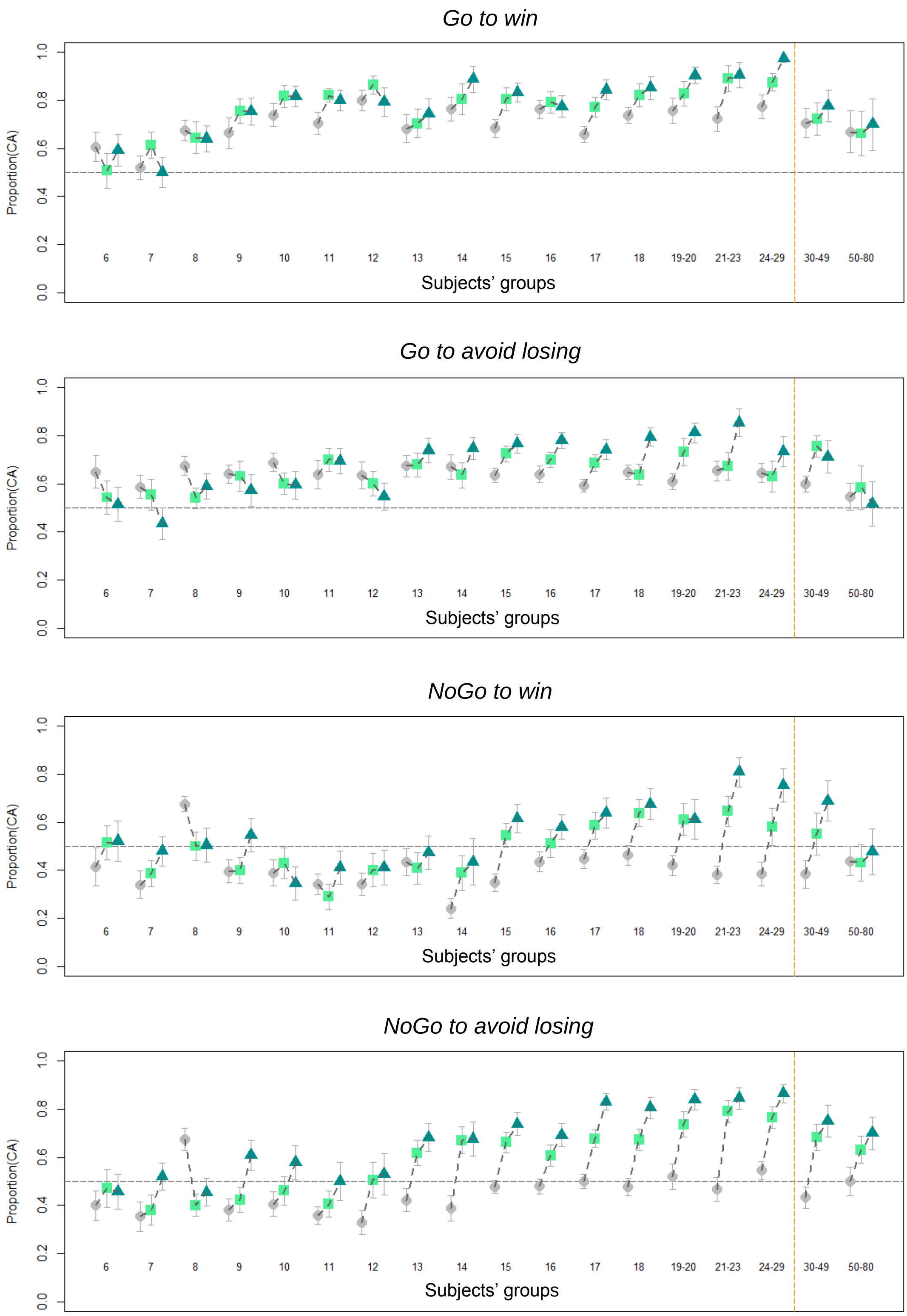


Figure 15. The proportion(CA) of each condition for each subjects' group on average. From left to right and top to bottom: *Go to win*, *NoGo to win*, *Go to avoid losing*, *NoGo to avoid losing*, and *Neutral* conditions. The horizontal dashed-line represents pressing the space-bar key by chance.

The results obtained dictate that the learning of each condition was verified at different subjects' ages:

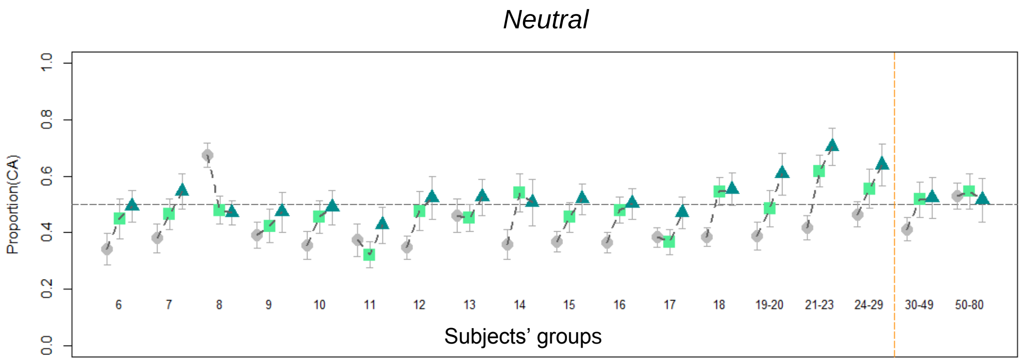
- *Go to win*: from 8 years-old;
- *Go to avoid losing*: from 8 until 30-49 years-old, inclusive;
- *NoGo to win*: from 15 years-old;
- *NoGo to avoid losing*: from 13 years-old.

In addition, the overall performance of each subjects' group can also be analysed by the evolution of the proportion of correct actions along the task (between the three blocks) for each condition (figures 16 and 17).



Legend:
1st Block
2nd Block
3rd Block

Figure 16. Averages of the proportions(CA) of each block per subjects' groups. From top to bottom: *Go to win*, *Go to avoid losing*, *NoGo to win* and *NoGo to avoid losing* conditions. The horizontal dashed-line represents pressing the space-bar key by chance and the vertical line separated the last two subjects' groups, whose data were not included in the mathematical models. Legend on the left.



Legend:
1st Block
2nd Block
3rd Block

Figure 17. Averages of the proportions(CA) of each block per subjects' groups of the *Neutral* condition. The horizontal dashed-line represents pressing the space-bar key by chance and the vertical line separated the last two subjects' groups, whose data were not included in the mathematical models. Legend on the left.

As it can be observed, each graph includes information about one of the five task condition, where each mark represents the average of the proportion(CA) for each block, being each cluster of three marks (the averages of the 3 blocks of the task) the overall performance of each subjects' group. Moreover, and as it was previously mentioned, learning a condition presupposes a gradual increasing of the values of the proportion(CA) across blocks, and that these values (at least the one corresponding to the last block) need to be higher than 0.5, which is the reached value when performing the task by chance. Therefore, the learning of the presented conditions was verified at different subjects' ages:

- *Go to win*: from 9 years-old;
- *Go to avoid losing*: from 11 years-old, excluding both 12 and 50-80 years-old;
- *NoGo to win*: from 15 years-old, excluding the 50-80 years-old;
- *NoGo to avoid losing*: from 9 years-old.

In both *Neutral* barplots (*Neutral* in figure 15 and figure17), it is possible to observe that the crescent tendency is very subtle, and better within the block's proportions' averages (figure 17), excepting for the 8 years-old subjects' group. It is worth noting that only the 19-20, 21-23 and 24-29 subjects' groups reached values above 0.6.

2.3.3. Model comparison

This section includes the obtained results of the developing process of the mathematical model that best fits the behavioural data. All investigated models included data from the task's performance of participants whose age-range between 6-30 years old (N = 382), being the model's data comprise by 1910 data points (382 x 5 conditions).

First of all, both linear and logit models presented by the previously 4 and 5 equations (page 50) were compared with the ones that did not include the subjects' age as independent variable, to verified if this variable was an important adding to the models. For this purpose, both BIC and AIC criteria were used, being the best model the one that possesses the lowest value of both criteria. The results are presented in table 7 and confirm that, in fact, both models benefit with the inclusion of the age variable.

Furthermore, all the remaining tested models used the participants' age (or an age variable transformation), the valence, action and neutral conditions' characteristics as additional independent variables of the model.

In order to simplify the writing of the models, each of these models were identified by one specific $y \sim x$ equation, being y the dependent variable, which is the proportion of correct actions (p) or a transformation of this variable, and the x represents the participants' age (age) variable, or its variable transformation. Table 8 presents both R^2 and Adj-R^2, the p-value and the Log-likelihood values of each investigated model.

Table 7. Both BIC and AIC criteria of linear and logit models, which included or not the age as an independent variable.

Models	BIC	AIC
Linear without the age variable	371.27	355.94
Linear with the age variable	**231.54**	**258.44**
Logit without the age variable	6950.94	6935.61
Logit with the age variable	**6810.64**	**6837.53**

Table 8. The R^2, Adj-R^2, p-value and log-likelihood values of tested models.

Tested models	R^2	Adj-R^2	p-value	Log-likelihood
$p \sim age$	0.2125	0.2088	2.8 e^{-92}	-72.26
$\boldsymbol{p \sim \log(age)}$	**0.2144**	**0.2107**	**2.9 e^{-93}**	**-69.93**
$p \sim \log(age) + age$	0.2128	0.2091	2.0 e^{-92}	-71.89
$p \sim \sqrt{age}$	0.2143	0.2106	3.2 e^{-93}	-70.05
$p \sim age^2$	0.2042	0.2005	4.9 e^{-88}	-82.22
$p \sim age^2 + age$	0.2046	0.2008	3.4 e^{-88}	-81.83
$\log(p) \sim age$	0.1612	0.1572	1.2 e^{-66}	-1740.35
$\log(p) \sim \log(age)$	0.1619	0.1580	5.2 e^{-67}	-1739.52
$\text{logit}(p) \sim age$	0.2082	0.2045	4.6 e^{-76}	-3281.36
$\text{logit}(p) \sim \log(age)$	0.2100	0.2062	5.3 e^{-81}	-3279.15
$\text{logit}(p) \sim \log(age) + age$	0.2085	0.2048	3.2 e^{-77}	-3281.00
$\text{logit}(p) \sim \sqrt{age}$	0.2100	0.2062	5.7 e^{-79}	-3279.23
$\text{logit}(p) \sim age^2$	0.2001	0.1963	6.1 e^{-67}	-3291.05
$\text{logit}(p) \sim age^2 + age$	0.2004	0.1967	4.2 e^{-67}	-3290.67

As it can be observed, the best model was the $p \sim \log(age)$ model, since it possessed the highest Log-likelihood value (highlighted in bold in table 8). The selected model might be represented by the following equation:

$$\begin{aligned} p \sim\ & \beta_0 + \beta_1 valence + \beta_2 neutral + \beta_3 action + \beta_4 valence \times action \\ & + \beta_5 \log(age) + \beta_6 valence \times \log(age) + \beta_7 neutral \times \log(age) \\ & + \beta_8 action \times \log(age) + \beta_9 valence \times action \times \log(age) \end{aligned} \tag{6}$$

In addition, table 9 includes the estimated parameters (β) of the selected model, as well as their standard errors and p-values. As it can be observed, neither β_1 nor β_4 parameters' estimation achieved statistical significance, demonstrating that it is still possible to improve the model.

Table 9. Estimated parameters of the selected model, including their estimate values (estimate), standard errors (Std. Error) and p-values (p-value).

Model's parameters	Estimate	Std. error	p-value
β_0	-0.250	0.097	**0.0096**
β_1	0.204	0.137	0.1354
β_2	0.499	0.137	**0.0003**
β_3	0.447	0.137	**0.0011**
β_4	-0.178	0.193	0.3580
β_5	0.340	0.036	**< 2.0 e^{-16}**
β_6	-0.118	0.051	**0.0221**
β_7	-0.245	0.051	**2.1 e^{-16}**
β_8	-0.156	0.051	**0.0024**
β_9	0.149	0.073	**0.0399**

2.3.3.1. Stepwise regression

To verify the importance of each explanatory variables included in the selected model (presented by the equation 6), a manual stepwise regression was made. The final results of this process are presented in the next table (table 10), where the best model was the one that has the lowest AIC and BIC values (highlighted in bold in table).

Table 10. The R^2, Adj-R^2, p-value, and both AIC and BIC values for the tested models.

	Tested models	R^2	Adj-R^2	p-value	AIC	BIC
	$p \sim \log(age)$ (equation 6)	0.2144	0.2107	2.86 e^{-93}	161.86	222.96
Excluded variables	$valence\ x\ action$	0.2141	0.2108	5.23 e^{-94}	160.71	216.25
	$valence\ x\ action$ $valence$	**0.2134**	**0.2106**	**1.18 e^{-94}**	**160.14**	**210.13**
	$valence\ x\ action$ $valence$ $valence\ x\ action\ x\ log(age)$	0.1830	0.1804	4.82 e^{-80}	230.84	275.28

Therefore, the best model was the one that excluded the interaction between valence and action ($valence \times action$), and the valence ($valence$) variables. In this case, this selected model was a simplification of the previously selected one, since it only included 7 (out of 9) explanatory variables and 8 (out of 10) parameters. In addition, this model proved to be better comparing either with the more complex and the simpler model. Therefore, for this moment on, this new model will be identified as the model_7 and it is presented by the next equation (equation 7):

$$
\begin{aligned}
p \sim\ & \beta_0 + \beta_1 neutral + \beta_2 action + \beta_3 \log(age) \\
& + \beta_4 valence \times \log(age) + \beta_5 neutral \times \log(age) \\
& + \beta_6 action \times \log(age) + \beta_7 valence \times action \times \log(age)
\end{aligned} \tag{7}
$$

Moreover, this model can be represented graphically by 5 lines in a plot of the proportion of correct actions versus subjects' age (figure 18), where each condition has its own mathematical expression, being this dependent on the code of their own explanatory variables. Therefore, table S2 (page 113) presents the model for each condition.

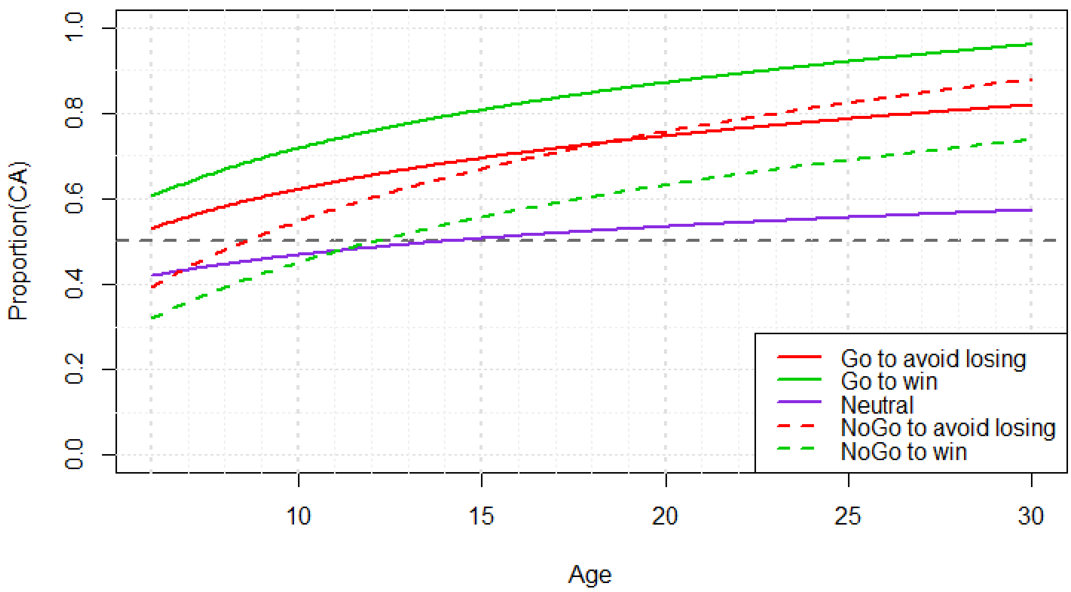


Figure 18. The predicted learning curves of the model_7 for all conditions. Each curve represents one condition as indicated by the legend presented on the right-bottom side of the graph. The horizontal dark-grey dashed-line represents pressing the space-bar key by chance – Proportion(CA) = 0.5.

In addition, figure S3 (page 114) offers a different view of the model, being the x-axis a logarithmic scale so that the different task conditions are represented as 5 straight lines allows to observe both slopes and intercepts of each condition.

That being said, the estimated parameters (β) of the model_7, as well as their standard errors and p-values are included in table 11. As it can be observed, all explanatory variables reached statistical significance (p-values under 0.05). In addition, the residual versus fitted values plots (figure S4 – page 114) supports that the model fitted well the behavioural data, as they were equally distributed around zero.

Finally, figure S5 (page 115) includes the 5 heat-maps representing the kernel density estimation (KDE) of the data used by the model. To better visualize the fit of the model, these graphs also included the averages, and respective SD, of the proportion(CA) of each subjects' groups (black points and respective error bars), and the prediction given by the model (black lines).

Table 11. Estimated parameters of the model_7, including their estimated values (Estimate), standard errors (Std. Error) and p-values (p-value).

Model's parameters	Estimate	Std. error	p-value
β_0	-0.148	0.068	0.0301
β_1	0.397	0.118	0.0008
β_2	0.358	0.097	0.0002
β_3	0.302	0.026	< 2.0 e^{-16}
β_4	-0.042	0.007	1.5 e^{-9}
β_5	-0.206	0.045	4.0 e^{-6}
β_6	-0.123	0.037	0.0008
β_7	0.083	0.010	< 2.0 e^{-16}

2.3.3.2. Contrasts

Taking into account the graphic representation of the model (figure 18), the asymptotes of each condition offered a good measure of learning. This measure was chosen due to the fact that slopes of the well-learned conditions were very small and, since they were directly related with the condition intercepts, they did not truly represented the learning process. As previously mentioned, both slopes and intercepts can be observed in both table S6 and figure S3 (pages 116 and 114, respectively).

Therefore, the greater the asymptotic value, which was computed using age = 30, the greater the learning of that specific condition. The following table (table 12) resumes the results of the computed asymptotic values for each task condition. As it can be observed, the highest value was reached for the *Go to win* condition (0.96), followed by the one for the *NoGo to avoid losing* condition (0.88), which are both considered congruent conditions. The next highest values were those obtained for the incongruent conditions: first, the *Go to avoid losing* (0.82), and, then, the *NoGo to win* condition (0.74). Finally, the learning of the *Neutral* condition was represented by the asymptotic value of 0.57. Noting that all asymptotic values were higher than 0.5, which should be the value obtained when subjects performed the task by chance.

Moreover, the contrasts were also used to analyse differences between congruent versus incongruent conditions, *Go* versus *NoGo* conditions, and win versus avoid losing conditions.

Table 13 includes the contrast weights for all of those aspects, which were computed in two steps: 1) by summing the contrast weights of the respective considered conditions, which are presented in the "Contrast weights" column (e.g.

Table 12. The contrasts weights and their respective values of each condition.

Conditions	Contrasts weights	Values
Go to avoid losing	$\beta_0 + \beta_2 + \beta_3 + \beta_6$	0.82
Go to win	$\beta_0 + \beta_2 + \beta_3 + \beta_4 + \beta_6 + \beta_7$	0.96
Neutral	$\beta_0 + \beta_1 + \beta_3 + \beta_5$	0.57
NoGo to avoid losing	$\beta_0 + \beta_3$	0.88
NoGo to win	$\beta_0 + \beta_3 + \beta_4$	0.74

the contrast weights of the congruent conditions were computed by summing the Go to win and the *NoGo* to avoid losing contrast weights conditions); 2) by subtracting the computed contrast weights to find potential differences between the two aspects, which are presented in the “Contrast results” column (e.g. Congruent – Incongruent). The last columns present both value and p-value of each analysed differences.

As it can be observed, these results indicate that the learning of the congruent conditions was 0.283 (p-value ~ 0) more effective than the learning of the incongruent conditions. Similar, the learning of the *Go* conditions was 0.163 (p-value = 0.008) more effective than the learning of the *NoGo* conditions. On the other hand, there was no statistical significant difference between the learning of both win and avoid losing conditions (p-value = 0.999).

Table 13. Both values and p-values of the contrast results of the differences between congruent and incongruent, go and *NoGo*, and win and avoid losing conditions.

	Conditions	Contrast weights	Contrast results	Values	p-values
Congruent	*Go to win*	$2x\beta_0 + \beta_2 + 2x\beta_3 + \beta_4 + \beta_6 + \beta_7$	β_7	0.28	~ 0
	NoGo to avoid losing				
Incongruent	*Go to avoid losing*	$2x\beta_0 + \beta_2 + 2x\beta_3 + \beta_4 + \beta_6$			
	NoGo to win				
Go	*Go to win*	$2x\beta_0 + 2x\beta_2 + 2x\beta_3 + \beta_4 + 2x\beta_6 + \beta_7$	$2x\beta_2 + 2x\beta_6 + \beta_7$	0.16	0.008
	Go to avoid losing				
NoGo	*NoGo to win*	$2x\beta_0 + 2x\beta_3 + \beta_4$			
	NoGo to avoid losing				
Win	*Go to win*	$2x\beta_0 + \beta_2 + 2x\beta_3 + 2x\beta_4 + \beta_6 + \beta_7$	$2x\beta_4 + \beta_7$	-4.8 e^{-5}	0.999
	NoGo to win				
Avoid losing	*Go to avoid losing*	$2x\beta_0 + \beta_2 + 2x\beta_3 + \beta_6$			
	NoGo to avoid losing				

3. Questionnaires' results

This section presents the statistical results of the data collected through the second part of the questionnaire, which included questions about the task performance: the self-evaluation of the subjects' attention during the task execution, the amount of different images, the images' beauty classification, and both best action and the points received, for both actions, for each image. The analysed questionnaires enclosed information of all participants (N = 413).

3.1. Attention during the task performance

The results of the subjects' self-evaluation concerning their attention during their task performance are presented in figure 19. It includes the relative frequencies of the "yes/no" subjects' responses to the 18th question of the questionnaire. Each bar, one for each group, includes the both subjects' answers "Yes, I was paying attention during the entire game" or "Yes" to simplify (represented by orange sections) and "No, I was not paying attention during the entire game", or "No" (represented by blue sections). The last bar (identified by "Total") includes the results of the entire sample population.

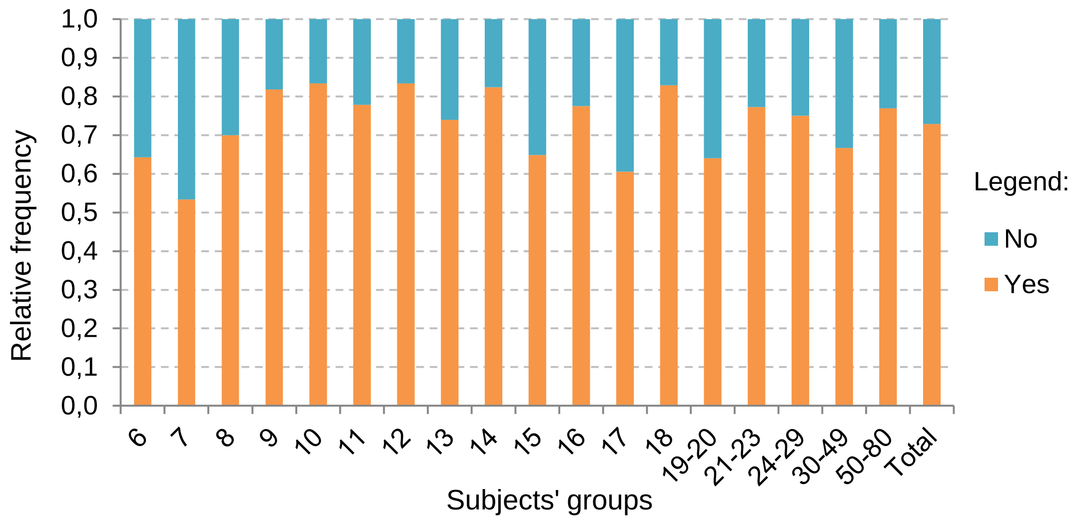


Figure 19. Relative frequencies of the subjects' answers to 18th question of the questionnaire. Legend on the right side of the graph.

In addition, table S7 (page 116) includes the relative frequencies presented in the graph. As it can be observed, the majority of the subjects' responses was "Yes" in all subjects' groups, being on average 0.73 and 0.27 for "Yes" and "No", respectively.

Furthermore, and as it was mentioned previously, when the subjects' response were "No", they had the opportunity of specified in which part(s) of the task (task's blocks or the entire task) they didn't pay attention, by answering to the 19th question of the questionnaire.

Similarly to the previously graph, figure 20 shows the relative frequencies of the 4 possible subjects' responses of all subjects' groups: "1st block" (dark red), "2nd block" (purple), "3rd block" (blue), and "Entire task" (orange). The last bar (also identified by "Total") represents the results obtained for all subjects. Table S8 (page 117) includes the relative frequencies presented in the graph.

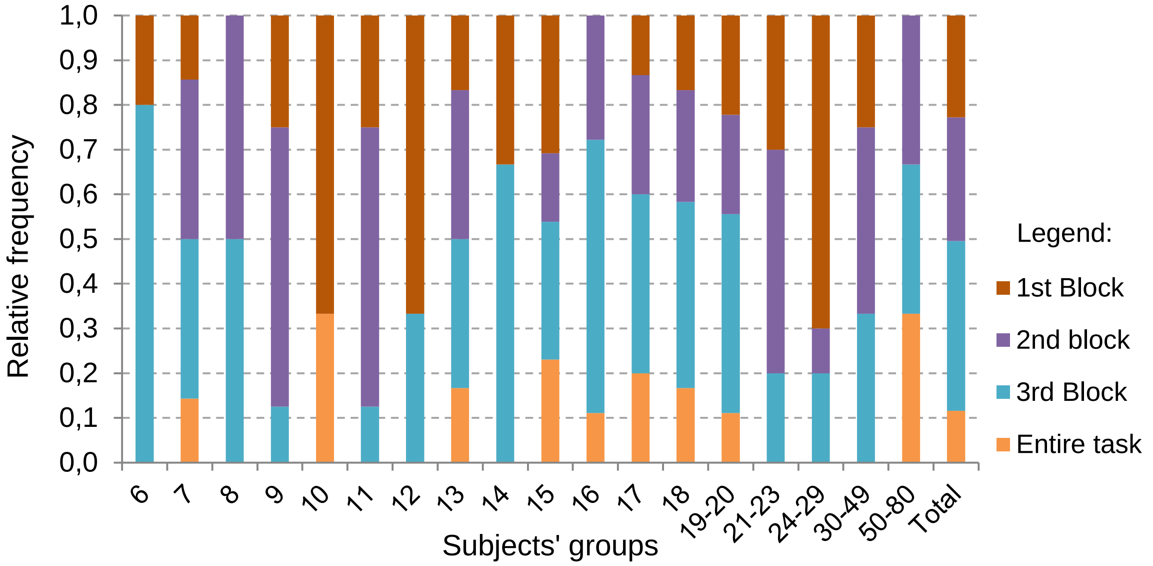
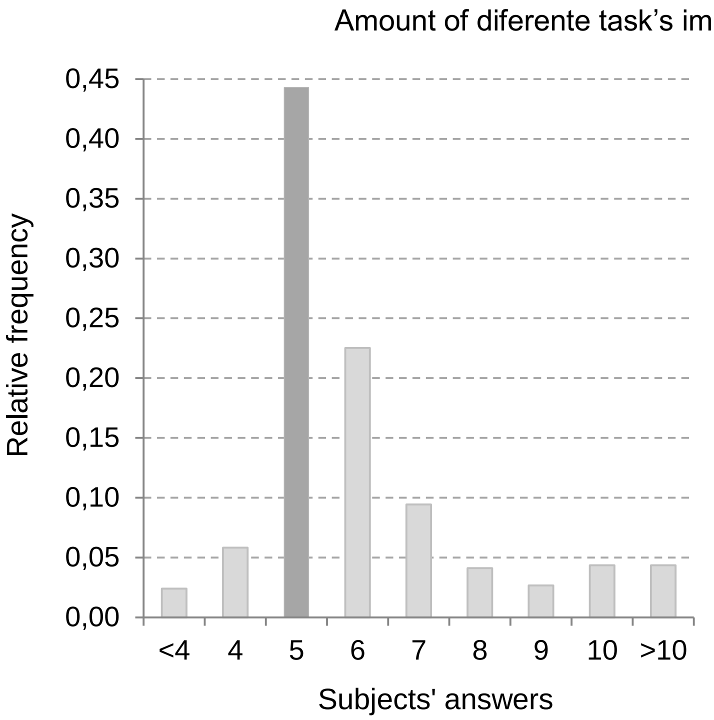


Figure 20. Relative frequencies of the 4 possible subjects' answers to 19th question of the questionnaire. Legend on the right side of the graph.

3.2. Amount of different task's images

The results of the analysis of the 20th question of the questionnaire are presented in both graph and table presented in figure 21.



Subjects' answers	Relative frequency
< 4	0.024
4	0.058
5	0.443
6	0.225
7	0.094
8	0.041
9	0.027
10	0.044
> 10	0.044

Figure 21. Relative frequencies of the subjects' answers to the 20th question of the questionnaire. Both graph (on the left) and table (on the right) share the same information, being highlighted the correct answer ("5").

As it can be observed, the highest frequency (44.3%) corresponds to "5", which was the correct answer, and the second most given response (22.5%) was "6". In fact, 47.5% of the subjects overestimated the amount of different images presented in the task, and only 8.2% under estimated that. In order to simplify the presentation of these results, some responses were clustered as <4 and >10. In addition, table S9 (page 117) includes the absolute frequency of all given answers.

3.3. Images' beauty classification

The results of the questions regarding the beauty of the images (21st-26th questions) were analysed firstly to verify if there was a clearly preference among the presented images of the task, and then to verify if the conditions behind them influenced the subjects' preferences.

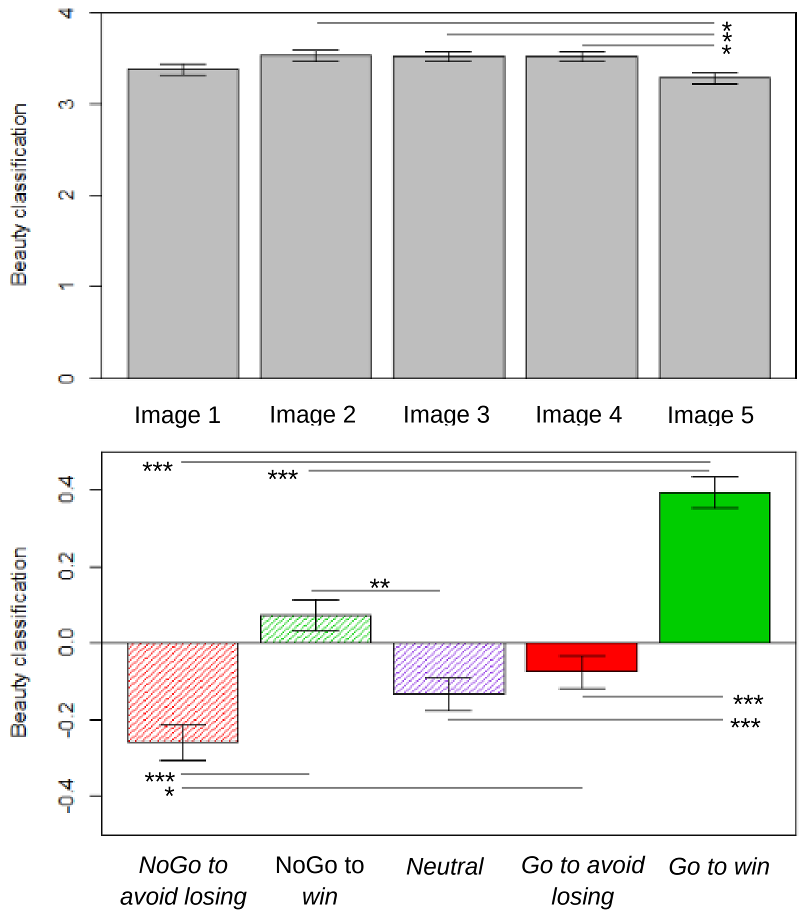


Figure 22. Beauty classification of each image (first graph), and of each condition (second graph). These results corresponded to the average and respective SD of the 21st-26th questions of the questionnaire.
F-test significance: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

The top graph of figure 22 includes the attributed rating (coded between 0 and 4 – 5 levels) concerning the beauty of each image. This analysis was independent of the conditions behind each image. The criteria for the nomination of each image was their appearance's order on the questionnaire.

Globally, subjects liked the images included in the task, classifying all images above 3 on average, which represents the "more or less beautiful" option. After running the ANOVA analysis and perform the Tukey post-hoc test, the differences between figure 5 and figure 2, 3, and 4 where statistically significant (p-value = 0.031, 0.039, and 0.046, respectively).

On the other hand, the second graph in figure 22 shows a completely different scenario. After matching each image with its correspondent condition, which varies between subjects, the z-scores were computed to normalize the subjects' classifications. Then, these values were used to calculate the averages of the beauty classification per condition.

As it can be observed, on average, both *avoid losing* and *Neutral* conditions, were the less preferred when compared with both *win*. Moreover, the congruent conditions show the two extremes classification: the *NoGo to avoid losing* condition received the most negative classification and, on contrary, the *Go to win* condition received the most positive classification.

Similar to last, after running the ANOVA analysis and perform the Tukey post-hoc test, some of the differences between conditions achieved the statistically significant value. The p-value between *Go to win* condition and all other were near zero, and the same happen to the differences between *NoGo to win* and *NoGo to avoid losing* conditions. In addition, the p-value between *NoGo to avoid losing* and *Go to avoid losing* conditions was 0.023, and it was 0.007 between the *Nogo to win* and *Neutral* conditions.

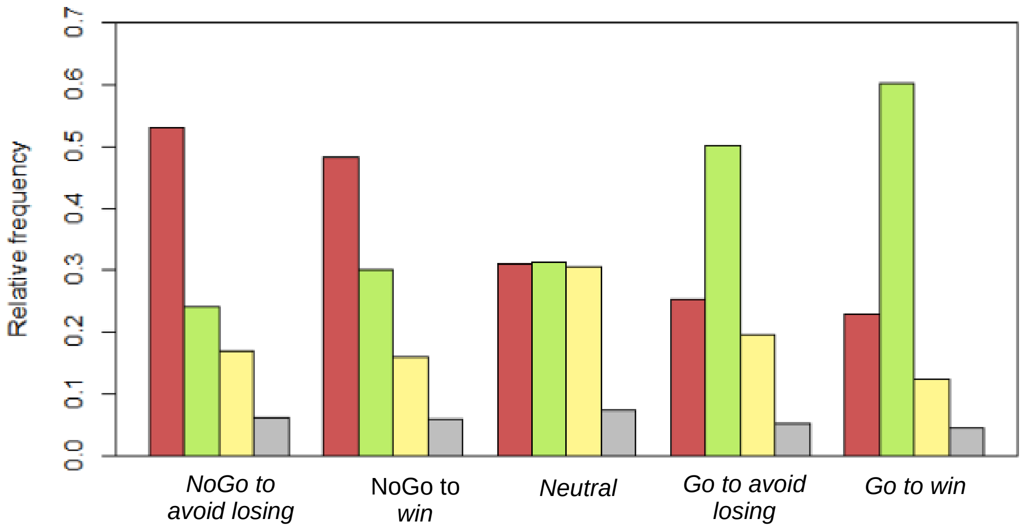
3.4. Best action and received points for each condition

Similarly, both best action and received points questions for each image were analysed according with the images' conditions.

Figure 23 summarizes the relative frequency of the 4 possible answers to the questions about which was the best action to execute given a specific image, being the data analysed regarding the condition behind each image. In this case, the questions relative to this data are the following: 26^{th}, 29^{th}, 32^{nd}, 35^{th}, and 38^{th}.

As it can be observed, the results indicate that subjects choose relatively more the option "*NoGo* action" to both *NoGo* conditions (*NoGo to avoid losing* and *NoGo to win*); and the option "*Go* action" to both *Go* conditions (*Go to avoid losing* and *Go to win*).

The *Neutral* condition received an equilibrated number of responses between "*NoGo* action", "*Go* action" and "It does not matter the action". Globally, subjects responded "I don't know / I don't remember" to less than 10% in all conditions.



Legend:

- *NoGo* action
- *Go* action
- It does not matter the action
- I don't know / I don't remember

Figure 233. Relative frequencies of the 4 possible subjects' answers to 27th, 30th, 33rd, 36th, 39th questions of the questionnaire. Legend on the left.

In addition, the received points for each condition depended on the action executed. Therefore, the results of both questions per condition included in the questionnaire were showed side by side on each of the graphs presented in figure 24. This figure included one graph for each task condition. To easily compare the results with the correct answers for these questions, table 14 includes the correct answers, which depended on which action executed, for each condition. Furthermore, the highlighted information indicates the cases where the most frequent given answer matched the correct one.

Table 14. Correct answers to the questions regarding which points were usually displayed on the screen (both actions)

Condition \ Action	***Go***	***NoGo***
Go to win	**+1**	**0**
NoGo to win	**0**	+1
Go to avoid losing	**0**	-1
NoGo to avoid losing	-1	**0**
Neutral	**0**	**0**

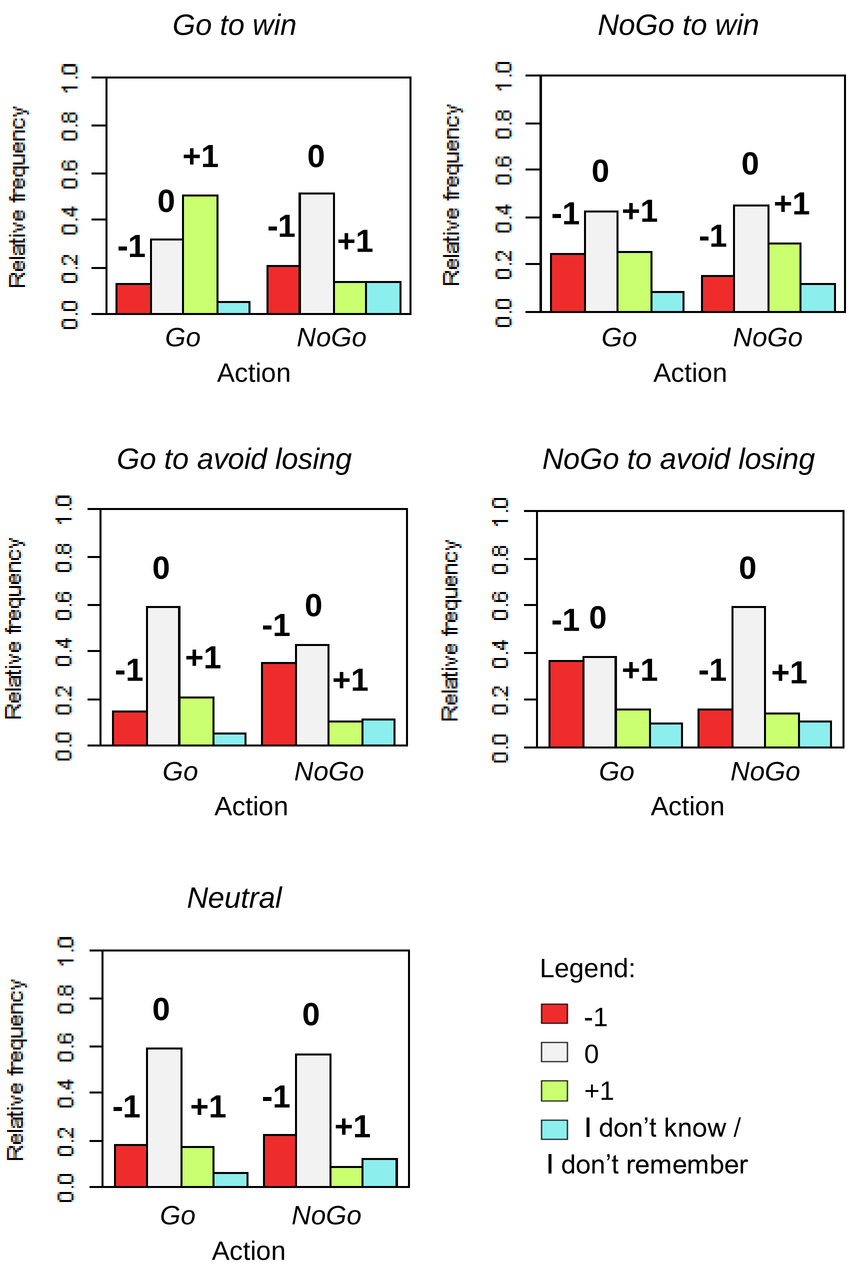


Figure 244. Received points for each condition. From left to right and top to bottom: *Go to win, NoGo to win*, *Go to avoid losing*, *NoGo to avoid losing*, and *Neutral* conditions; and the respective legend.

As it can be observed, the subjects' majority correctly responded to the both questions regarding the *Go to win* and *Neutral* conditions.

Regarding the *NoGo to win* condition, the results are very similar to the ones obtained for the *Neutral* condition, excepting the higher values for the "0" option in *Neutral* condition (0.6 versus 0.4), and the higher value for the "+1" response to the *NoGo* action in the *NoGo to win* condition. The latter may be a consequence of the subjects' learning, since this condition was the last to be learned so that, on average, only subjects older than 15 years old appeared to be capable of learning it.

The obtained results for the *Go to avoid losing* and *NoGo to avoid losing* conditions appeared to be symmetrical. In this case, the subjects' majority correctly responded to which correct action the "0" is the correct answer (*NoGo* action for the *Go to avoid losing condition*; and *Go* action for the *NoGo to win* condition). In addition, it is clear that some subjects learned that the other action had a negative value, being the correct answer "-1", this results also suggest that some subjects may not realize this negative value because they may not try that action, preferring execute the save action that always gave "0", enough times to have to have confidence in their answer.

Discussion

This chapter discusses the previously described results and it is organized in three parts: 1) some general considerations about the study in global; 2) a more detailed discussion of the results regarding the influence of age in the RL process and motor biases, through the analysis of the data obtained with the task performance; and 3) a more detailed discussion of the results regarding the consciousness level of the people execution of the task, through the analysis of the data obtained with the application of the final questionnaire.

1. General considerations

As mentioned previously, there are a future interest in applying this tested Go/NoGo task to study the RL process in people with some neurodevelopmental disorders, such as TS, OCD and ADHD. Since these disorders affect more children and teenagers than adults, it became essential to test the Go/NoGo task reliability when applied to younger people, once this task had already proved to be adequate to the RL study when applied to healthy adults (Dias, 2014).

With this in mind, this study started with the development of a successful recruitment process, especially for the youngest participants from schools. Furthermore, the sample size (N = 419), coupled with a balanced subject distribution within the considered age-range (from 6 to 80 years old), and gender (51.6% males), makes this a high quality data-base for studying RL in humans using this particular Go/NoGo task, which can be used further as a control population.

Testing this task in such a large well-characterized sample gave the opportunity to better understand the influence of age in the RL process, especially during the first two decades of human lifespan, which are considered the most critical to neuronal development (Shephard, Jackson, & Groom, 2014). Despite the fact that the Go/NoGo task intended to measure the RL process, it was also able to capture several differences within subjects' performances, which not only depended on their ages but also, more interestingly, on the task conditions.

In addition, the final questionnaire allowed collecting some additional demographic data, which can be used in a future analyses, and verifying the consciousness level of the subjects' performance.

2. Influence of age in both RL and motor biases

This study included 3 major measures to evaluate the RL process. Firstly, the learning curves (figures 11-13), which used the *Go* probabilities versus the trial number of the task, confirmed that the learning process happened gradually during the task execution, which depended on both task conditions and subjects' ages. Secondly, the proportion of correct actions using the last 20 trials (the last 2 blocks) of each condition allowed to easily compare the learning rate of each age group (figure 15). Finally, the proportion of correct actions of each condition, per block, and for each age group (figures 16-17) complemented the previous analysis. This latter analysis merged the two former analyses, since it enabled the comparison between the RL process along the task (by analysing the evolution of the resulting pattern of each cluster of the 3 block's proportions) and across subjects' ages (by analysing the tendency evolution between the clusters of the 3 blocks' proportions). Furthermore, the learning curves and the proportion between blocks also captured the learning speed of each condition, which also depended on the subjects' ages.

Once the presented results in figures 16 and 17 offered a more complete vision of the subjects' capability of learning each task conditions, the following discussion will be based on them. It is worth noting that these results included the performance analyses of 413 subjects, since 6 subjects were excluded due to displaying abnormal task solving capabilities, whose age-range allowed to split the sample in two major groups: between 6 and 30 years old (N = 382, 92.4%), and between 30 and 80 years old (N = 31, 7.6%).

Although the RL process occurred gradually during the solving of the Go/NoGo task, learning each condition was not only depended on subjects' ages by also a cumulative process. In other words, it was possible identify a specific order, within subjects' ages, in which the conditions began to be learned. It was extremely difficult to interpret these results, as it implicated to take into consideration the following aspects: 1) the motor biases identified by the *Neutral* condition; 2) the pattern shown between the averages of the 3 blocks, per age group; and, finally, 3) the values of those averages.

The first point was very important in a way that prevented false conclusions when subjects showed a dual learning pattern (only *Go* or *NoGo* learning), which might be highly dependent on their motor biases. In practice, this was considered by crossing the results of the learning of both *Go* and both *NoGo* conditions, with the respectively motor biases cached by the *Neutral* condition. When the subjects "learned" only both *Go* or both *NoGo* conditions and presented the same motor bias, *Go* or *NoGo* respectively, they cannot be considered real "learners". In this case, their learning is masked by their own motor tendency of responding *Go* or *NoGo* to all stimuli, without discriminating them.

Both second and third points, clearly demonstrated the importance of a careful analysis of these data. If on one side, pressing the space-bar key above 50% of the times does not mean a sustainable learning process, which was defined by the crescent values along blocks; on the other side, despite the evident learning across the 3 blocks, these average values (at least the one for the last block) might not reach the cut-off of 0.5. In the last situation, a conclusion that learning did not occurred seems precipitated. In fact, it is possible that some ages may require an additional period for learning specific conditions and thus, in a longer task version these values could possibly reach higher values.

Hereupon, subjects appeared to be capable of learning both *Go to win* and *NoGo to avoid losing* conditions (both congruent) at the age of 9-10, being the highest value reached by the former. Then, the learning of both *Go to avoid losing* and *NoGo to win* conditions (both incongruent) appeared to occur around 11-13 and 15 years old, respectively. As previously explained, the *NoGo to win* condition clearly represented the case where, despite the gradual increase of the learning pattern between the blocks' values started to be noticed earlier (around 11 years old), only later (at 15 years old) the performance exceeds the cut-off of 0.5. In this case, with a longer task version, subjects within this age-range (11-15 years old) might reach higher performance values.

Regarding the possible existence of different motor biases across ages, the analysis of the graph of the *Neutral* condition (figure 17) partially excluded them. As it can be observed, although not completely independent of the feedback, the 1st block value, per age group, showed the initial motor tendency for responding to a certain stimulus. Globally, the majority of the subject groups showed an initial *Go* bias, pressing the space-bar between 70 to 50% of the time during the 1st block. On the other hand, both 8 and 50-80 years old groups presented an initial *NoGo* bias, pressing the space-bar key around, respectively, 30% and 45% of the times during the 1st block. The latter may be a reflection of the subjects' unfamiliarity of using a computer and, in fact, three subjects (out of 4) confessed that the execution of this task was their first contact with a computer. The case of the 8 years old group will be discussed later.

Interestingly, the 8 year old group presented the highest *NoGo* bias in the first block for both *Neutral* and *NoGo* conditions, which contrasted with the also higher *Go* values in the first block of both *Go* conditions. This might suggest that, on average, the 8 years old children may have started to learn both *Go* and *NoGo* conditions during the 1st block, and then may lost the focus or get boring.

In addition, as it can be observed, since both 6 and 7 years old subjects' groups showed an identical *NoGo* performance pattern, meaning that subjects gradually stopped pressing the space-bar key (from 0.6 to 0.5-0.4, approximately). In this case, the subjects' performance cannot be seen as an active learning process, since it revealed to be a dual learning influenced by their *NoGo* bias.

Surprisingly, on average, the 12 years old group showed an inferior performance in both *Go* conditions when compared with the 11 years old group, especially in the *Go to avoid losing* condition. On the other hand, both 11 and 12 years old groups displayed an inferior performance in the *NoGo to avoid losing* condition when compared to the increasing tendency observed in both preceding and succeeding groups. On one hand, these inconsistencies may be due to a specific effect of the age, being these ages known to be the beginning of the adolescence, which are characterized by several psychological and physical changes (Casey, Giedd, & Thomas, 2000). On the other hand, it can be due to a specific school influence, since these subjects were mainly recruited from the same specific school (EB23C). In the future, and to better understand these results, it is essential to collect data from subjects within the same age-range from a different environment.

Furthermore, since this study focused particularly in the recruitment of young subjects, only a small amount of older subjects (N = 31, age-range from 30-80) were included, being their age distribution not ideal. In short, these results are bound not be truly representative regarding this age-range. Despite all that, it is possible to verify that the performance of these older subjects declined when compared to the previous groups. In fact, the oldest subjects, whose age-range was 50-80 (the last group), only learned the congruent conditions (*Go to win* and *NoGo to avoid losing*), thus showing a behavioural learning very similar to the subjects from the 9 years old group. In contrast, the highest performance values were achieved by the young adults, whose age-range was between 21 and 30 years old.

That being said, it is possible that the RL process, which was measured by the behavioural data collected through the application of this Go/NoGo task, may present an inverted-U shape relation with the subjects' ages (Bäckman, *et al.*, 2006). Furthermore, the obtained total scores (figure 14) also presented an inverted-U shape across age. Since this measure revealed to be globally consistent with the learning across ageing, especial between 13 and 18 years old groups, it may be adopted in the future as a fast index of the RL process.

In summary, this study proved that the capability of solving this RL task gradually increases during the neurodevelopment of children and teenagers, achieving its peak in young adults and then starts to decline during ageing.

2.1. Modelling the behavioural data

In order to increase the understanding of the mathematical relation between the behavioural data collected through the Go/NoGo task and the subjects' ages, this study also included the development of a mathematical model.

This model aimed to fit the behavioural data, which in this case were the individual five proportions of correct actions (represented by *p*), using the last 20 trials of each task condition, versus the respective subjects' ages (represented by *age*), and only included data for subjects younger than 30 years old.

After testing several models, the final model, denominated as model_7 (equation 7 – page 71), included 7 explanatory variables, all statistically significant (p-value < 0.05, table 11), where the age variable was transformed into the logarithm of age. This means that the Go/NoGo task performance followed a logarithmic shape along the subjects' ages (figure 18).

In addition, despite the obtained low-values of both R^2 and Adj-R^2 values (0.2134 and 0.2106, respectively), the model reached a higher statistical significance (p-value = 1.18 e^{-94}, table 10). Furthermore, figure S5 (page 115) offered a complete vision of how well the model fits the behavioural data.

Globally, the predicted data from the model evidence the same temporal sequence of the described learning. Despite that, the model overestimated the learning process in around 2 years earlier than the previously discussed. This can be also verified through the analysis of figure 15, which presents the averages of the proportions of the correct actions using the last 20 trials of each condition for each subjects' group. One possible solution to this misfit is to consider the trial number as an additional independent variable.

Apart from that, the results obtained using the contrasts were consistent with the ones already described: the learning of the congruent conditions was 0.28 greater (p-value ~ 0) than the learning of the incongruent conditions; and the learning of both *Go* conditions was also 0.16 greater (p-value = 0.0008) than the learning of both *NoGo* conditions. In addition, no significant differences were found between learning the win conditions and the avoid losing conditions (p-value = 0.999).

3. Evaluation of the consciousness level of the task performance

In general, the subjects' self-evaluation about their attention during their own task solving was positive, whereas 73% considered to pay attention during the entire task – figure 19 (and table S7). On one hand, subjects included in the 7 years old group were the ones that felt more difficulties to pay attention to the task (47%). On the other hand, only 17% of the subjects included in the 10, 12 and 18 years old groups felt some difficulties. In addition, these subjects had also the opportunity to specify in which part(s) of the task they considered to have lost their attention – figure 20 (and table S8). In general, the "3rd block" was the most chosen task part (38%) where subjects felt paying less attention, followed by the "2nd block" (28%), "1st block" (23%), and "entire task" (12%) answers. This results may be a consequence of the sustained mental demand needed for the correct task solving.

Furthermore, the analysis of the amount of different images displayed along the task also offered a global view regarding the task performance (figure 21). In general, 44.3% of the subjects correctly answered "5" images, 22.5% and 9.4 % answered, respectively, "6" and "7" images. In fact, 47.5% of all subjects gave a response above 5, whereas only 8.2% gave an inferior response. This overestimation may reflect the subjects' perception that this task was more complex than it actually was, or they may recall some memories of the images previously displayed during both demonstration and training phases, which were different from the ones used in the testing phase.

Globally, the analysis of the rating concerning the beauty of the used images showed a surprisingly consistent relation between the attributed scores and the conditions behind each image (second graph of figure 22). Globally, and despite the warning saying that this classification should be independent of feedback received during the task solving, people positively rated both *win* conditions, and negatively rated both *avoid losing* and *Neutral* conditions. The congruent conditions received the positive highest (*Go to win* condition) and negative lowest (*NoGo to avoid losing*) scores. On the other hand, since image 5 showed statistically significant differences between images 2, 3, and 4 (first graph of figure 22), an important further analyses will be verify if this result may be due to the condition/image distribution, since it was not equilibrated.

Finally, the last questions of the questionnaire were designed to analyse the subjects' consciousness level of their own task solving. Taking into account their own task, for each displayed image, subjects had to indicate what was the best action to perform during the game (figure 23), and which points they usually received after executing the *Go* and the *NoGo* action (figure 24).

As it can be observed in figure 23, the majority of the answers were given correctly to the questions regarding which action was the best for each image, wherein this analysis took into account its condition: "*Go* action" to both *Go* conditions and "*NoGo* action" to both *NoGo* conditions. Despite the considered the most correct answer for the Neutral condition be "It does not matter the action", the subjects' answers showed an equilibrated distribution also between "*Go* action" and "*NoGo* action" options. Moreover, only less than 10% subjects answered "I don't know / I don't remember". Globally, these results demonstrated that subjects were aware of their performance few minutes after solving the task. In the future, it will be interesting to see how these results can be correlated with the conditions' learning level, which obvious will be also influenced by the subjects' ages.

Regarding the results about the received points (figure 24), the answers showed a correlation between conditions' learning and subjects' confidence level on their answers. In this case, a confident answer is the one that stands out from the rest.

That being said, the subjects' majority correctly responded to the both questions regarding the *Go to win* and *Neutral* conditions; and to the "safe" action (being the feedback always "0") in both *Go* and *NoGo to avoid losing* conditions. In fact, although it was clear that some subjects learned the negative value of the other action, being the correct answer "-1", this results also suggest that some subjects may not realize this negative value because they may not try that action enough times to have to have confidence in their answer preferring execute the save action instead. Regarding the *NoGo to win* condition, the results may be a consequence of the subjects' learning, since this condition was the last to be learned so that, on average, only subjects older than 15 years old appeared to be capable of learning it.

In summary, both results to the questions of what was the best action and which points were more often received, for each condition, indicates that, on average, people remember their task solving few minutes after doing it. Despite that fact, further analysis will help to understand if this learning process also requires the involvement of the working memory and, possible, the hippocampal memory. One possible way to test these aspects may be present these same questions a longer period after solving the task.

To finish, our data indicates that subjects, on average, solved this task in a consciousness way, since they were not able of recall the memories about their performance (Kandel, Schwartz, & Jessell, 2000). In addition, similarly to what has been previously discussed, it will be important to do further analysis to correlate the task's learning level, the subjects' ages and, also, their capability of give correct answers to these questions.

Conclusions and Future work

This last chapter encloses the main conclusion obtained during this study as well as some of the already planned future work

1. Conclusions

The main goal of this study was to have a better understanding of the influence of age in the reinforcement learning process in humans, by applying a new probabilistic Go/NoGo cognitive task followed by a final questionnaire. Having this in mind, we established a demanding recruitment process in order to obtain not only the largest sample of healthy people that was possible, but also to try to get a well distributed sample concerning their ages. In addition, and due to further interests in applying this same task in younger people with neurodevelopment disorders (such as TS, OCD and ADHD), our focus was directed to the recruitment of mostly young subjects. Therefore, this process resulted in a sample of 419 healthy subjects, who aged from 6 to 80 years old, being 92,4% younger than 30 years old.

The applied novel probabilistic RL Go/NoGo task showed promising results to a further studies of the RL study in humans. To briefly summarize, the task included 5 conditions, being a probabilistic feedback displayed regarding the executed action. In two conditions (*Go to win* and *NoGo to win*) the aim was to score points, in other two conditions (*Go to avoid losing* and *NoGo to avoid losing*) the aim was to avoid losing points, and finally, in the last condition (*Neutral*), the feedback was always zero independently of which performed action.

Globally, the task design, proved to be sensitive to both positive and negative learnings, being the neutral condition a good index of the subjects' motor biases. In addition, the task was not only well-succeeded to encompass differences during the task execution due to the influence of subjects' age, but it was also sensitive enough to individually characterize the learning process among different subjects, being possible to seem a unique performance profile for each subject.

Furthermore, all collected evidences pointed out to the existence of a unique learning sequence among the Go/NoGo task conditions, which depends on age. Our data suggested that learning from both conditions *Go to win* and *NoGo to avoid losing* begin around the age of 9, which is then followed by the learning of the *Go to avoid losing* condition, around the age of 11; and, finally, learning the *NoGo to win* condition seems to happen at 15 years old. In addition, the developed mathematical model, in spite of not being perfect, also proved to be an important tool to help understanding the RL process in humans.

Finally, the questionnaire offered a great opportunity to access the subjects' consciousness level of the execution of the task. Our results indicates that subjects were aware of their own performance, which suggested that this may be considered a conscious process, possible using also other memories systems, such as the working and the hippocampal memories, to complement the automatic one. Surprisingly, the results of the images' beauty classification revealed an unconsciousness preference for the images with the *win* conditions behind it than the ones with *the avoid losing* or *Neutral* conditions.

Globally, this study was an important step in understanding the RL process in humans. Moreover, this study also gave the opportunity to demonstrate the reliability of this Go/NoGo task application.

2. Future work

Due to both high quantity and quality of the collected data, which included the data from the task performances and from the questionnaires answers, it is still possible to make several interesting analyses, some of which are already planned.

Firstly, as expected, we want to optimize the mathematical model by adding the trial number as an additional independent variable and substituting the actual dependent variable by the actions executed by the subjects. In addition, and since this task may have a clinical use in the future, the model will add an important value to it. Ideally, this optimal model will be able to predict the population's normative behavioural data, which means that with the data from the task execution and the age of subjects, it will be possible to infer in which learning percentiles these subjects belong to.

Furthermore, due to the design of the task, namely the use of random durations between some elements within the trials, this task is also ready to be used in a fMRI environment. In fact, at the moment we are starting to collect some preliminary data about the behavioural performance in a fMRI environment, where we expect that the dopaminergic activity measured on the basal-ganglia, in particular in the striatum, can be related to the occurrence of prediction errors, which are a consequence of using probabilistic feedback. In addition, some genetic analyses can complement both behavioural and fMRI data, especially the ones within the striatal dopaminergic genes (e.g. DRD1, DRD2, DAT).

In the future, we are also planning to develop some integrative studies, including both fMRI imaging and genetic analyses to the behavioural data obtained through the application of Go/NoGo task. In fact, the evaluation of the RL process in some neurodevelopment disorders might be the key to better characterize their psychological profiles, especially the involvement of possible dopaminergic disruption. All these evidences may lead to the discovery of new links between highly comorbidities disorders, such as the TS and both ADHD and OCD, or to increase the knowledge about some endophenotypes.

The development of new cognitive tasks, specialized for the analyses of determined neuronal pathways will change the way that clinicians look into psychological disorders. These new tools will offer higher analytical insight of these disorders that may lead to an improvement of both behavioural and pharmacologic therapies and, consequently, to a better quality of life for the affected people.

References

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly journal of experimental psychology*, *33*(2), 109-121.

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in neurosciences*, *12*(10), 366-375.

Alexander, R. D. (1974). The evolution of social behavior. *Annual review of ecology and systematics*, 325-383.

Annau, Z., & Kamin, L. J. (1961). The conditioned emotional response as a function of intensity of the US. *Journal of comparative and physiological psychology*, *54*(4), 428.

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S. C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: current status and future prospects. *Neuroscience & Biobehavioral Reviews*, *30*(6), 791-807.

Baldwin, R. L., Chelonis, J. J., Prunty, P. K., & Paule, M. G. (2012). The use of an incremental repeated acquisition task to assess learning in children. *Behavioural processes*, *91*(1), 103-114.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129-141.

Bayer, H. M., Lau, B., & Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *Journal of Neurophysiology*, *98*(3), 1428-1439.

Bogacz, R., & Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural computation*, *19*(2), 442-477.

Cachope, R., & Cheer, J. F. (2014). Local control of striatal dopamine release. *Frontiers in behavioral neuroscience*, *8*.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological psychology*, *54*(1), 241-257.

Chevalier, G., & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in neurosciences*, *13*(7), 277-280.

Cohen, M. X., & Frank, M. J. (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behavioural brain research*, *199*(1), 141-156.

Colwill, R. M., & Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of experimental psychology: animal behavior processes*, *11*(1), 120.

Crone, E. A., Jennings, J. R., & Van der Molen, M. W. (2004). Developmental change in feedback processing as reflected by phasic heart rate changes. *Developmental psychology*, *40*(6), 1228.

Darwin, C. (2009). *The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life*. W. F. Bynum (Ed.). AL Burt.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876-879.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology*, *18*(2), 185-196.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011, May). Gamification. using game-design elements in non-gaming contexts. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 2425-2428). ACM.

Dias, A. R. N. (2014). *Mechanistic characterization of reinforcement learning in healthy humans using a model-based approach.* Master dissertation, Instituto Superior Técnico, Universidade de Lisboa.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *308*(1135), 67-78.

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *The Journal of neuroscience*, *31*(16), 6188-6198.

Eppinger, B., Hämmerer, D., & Li, S. C. (2011). Neuromodulation of reward-based learning and decision making in human aging. *Annals of the New York Academy of Sciences*, *1235*(1), 1-17.

Eppinger, B., Schuck, N. W., Nystrom, L. E., & Cohen, J. D. (2013). Reduced striatal responses to reward prediction errors in older compared with younger adults. *The Journal of Neuroscience*, *33*(24), 9905-9912.

Frank, M. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Cognitive Neuroscience, Journal of*, *17*(1), 51-72.

Frank, M. J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Current opinion in neurobiology*, *21*(3), 381-386.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological review*, *113*(2), 300.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(2), 137-160.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007, b). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences*, *104*(41), 16311-16316.

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007, a). Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, *318*(5854), 1309-1312.

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940-1943.

Frank, M., & Fossella, J. (2010). Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology*, 36(1), 133-152.

Frank, M., & Hutchison, K. (2009). Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience*, 164(1), 131-140.

Frank, M., & O'Reilly, R. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci*, 120(3), 497-517.

Gleitman, H. (1995). *Psychology.* [Psicologia]. New York, NY: W. W. Norton & Company, Inc.

Gluck, M. A., Mercado, E., Myers, C. E. (2008). *Learning and Memory: From Brain to Behavior*. New York, NY: Worth Publishers.

Hämmerer, D., Li, S. C., Müller, V., & Lindenberger, U. (2011). Life span differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *Journal of Cognitive Neuroscience*, *23*(3), 579-592.

Hernández-López, S., Bargas, J., Surmeier, D. J., Reyes, A., & Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *The Journal of neuroscience*, *17*(9), 3334-3342.

Hernández-López, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., & Surmeier, D. J. (2000). D2 dopamine receptors in striatal medium spiny neurons reduce L-Type Ca2+ currents and excitability vía a novel PLCβ1–IP3–calcineurin-signaling cascade. *The Journal of Neuroscience*, *20*(24), 8987-8995.

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A theory. *Behavioral biology*, 12(4), 421-444.

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research Reviews*, *31*(1), 6-41.

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Powers, A., Mehta, N., & Casey, B. J. (2014). Adolescent-specific patterns of behavior and neural activity during social reinforcement learning. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 683-697.

Joshua, M., Adler, A., & Bergman, H. (2009). The dynamics of dopamine in control of motor behavior. *Current opinion in neurobiology*, *19*(6), 615-620.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. *Punishment and aversive behavior*, 279-296.

Kandel, E. R., Schwartz, J. H., Jessell, T. M. (2000). *Principles of Neural Science*. New York, NY: McGraw-Hill Companies.

Lee, D. (2013). Decision making: from neuroscience to psychiatry. *Neuron*, *78*(2), 233-248.

Ludvig, E. A., Bellemare, M. G., & Pearson, K. G. (2011). A primer on reinforcement learning in the brain: Psychological, computational, and neural perspectives. *Computational neuroscience for advancing artificial intelligence: Models, methods and applications*, 111-144.

Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(4), 343-364.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature neuroscience*, *14*(2), 154-162.

Maren, S., & Holt, W. (2000). The hippocampus and contextual memory retrieval in Pavlovian conditioning. *Behavioural brain research*, *110*(1), 97-108.

Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Progress in neurobiology*, *50*(4), 381-425.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. *Neurobiology of learning and memory*, 65-77.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of neuroscience*, *16*(5), 1936-1947.

Moustafa, A. A., Cohen, M. X., Sherman, S. J., & Frank, M. J. (2008). A role for dopamine in temporal decision making and reward maximization in parkinsonism. *The Journal of Neuroscience*, *28*(47), 12294-12304.

Nambu, A., Tokuno, H., & Takada, M. (2002). Functional significance of the cortico–subthalamo–pallidal 'hyperdirect'pathway. *Neuroscience research*, *43*(2), 111-117.

Nambu, A., Tokuno, H., Hamada, I., Kita, H., Imanishi, M., Akazawa, T., Ikeuchi, Y., & Hasegawa, N. (2000). Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *Journal of neurophysiology*, *84*(1), 289-300.

Nemoda, Z., Szekely, A., & Sasvari-Szekely, M. (2011). Psychopathological aspects of dopaminergic gene polymorphisms in adolescence and young adulthood. *Neuroscience & Biobehavioral Reviews*, *35*(8), 1665-1686.

Nicola, S. M., Surmeier, D. J., & Malenka, R. C. (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annual review of neuroscience*, *23*(1), 185-215.

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., & Pessiglione, M. (2009). Brain hemispheres selectively track the expected value of contralateral options. *The Journal of Neuroscience*, *29*(43), 13465-13472.

Pavlov, I. P. (1927). Conditioned reflexes. An Investigation of the physiological activity of the cerebral cortex. Oxford, England: Oxford University Press.

Pawlak, V., & Kerr, J. N. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *The Journal of Neuroscience*, *28*(10), 2435-2446.

Penny, W. D. (2012). Comparing dynamic causal models using AIC, BIC and free energy. *Neuroimage*, *59*(1), 319-330.

Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem?. *Neuroscience*, *89*(4), 1009-1023.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological review*, *74*(1), 71.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: current research and theory*.

Reynolds, J. N., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*(6851), 67-70.

Richfield, E. K., Penney, J. B., & Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D 1 and D 2 receptors in the rat central nervous system. *Neuroscience*, *30*(3), 767-777.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, *80*(1), 1-27.

Schultz, W. (2007). Behavioral dopamine signals. *Trends in neurosciences*, *30*(5), 203-210.

Scoville, W. B., & Milner, B. (1957*).* Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, 20(1), 11.

Seger, C. A., & Spiering, B. J. (2011). A critical review of habit learning and the Basal Ganglia. *Frontiers in systems neuroscience*, *5*.

Shephard, E., Jackson, G. M., & Groom, M. J. (2014). Learning and altering behaviours by reinforcement: Neurocognitive differences between children and adults. *Developmental cognitive neuroscience*, *7*, 94-105.

Skinner, B. F. (1938). The behavior of organisms: An experimental analysis.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory*, *82*(3), 171-177.

Strauss, G. P., Frank, M. J., Waltz, J. A., Kasanova, Z., Herbener, E. S., & Gold, J. M. (2011). Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biological psychiatry*, *69*(5), 424-431.

Surmeier, D. J., Plotkin, J., & Shen, W. (2009). Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. *Current opinion in neurobiology*, *19*(6), 621-628.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT press.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs: General and Applied*, *2*(4), i-109.

Thorndike, E. L. (1911). Animal intelligence: Experimental studies. Macmillan.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. Univ of California Press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.

Wimmer, G. E., Daw, N. D., & Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. *European Journal of Neuroscience*, *35*(7), 1092-1104.

Worbe, Y., Palminteri, S., Hartmann, A., Vidailhet, M., Lehéricy, S., & Pessiglione, M. (2011). Reinforcement learning and Gilles de la Tourette syndrome: dissociation of clinical phenotypes and pharmacological treatments. *Archives of general psychiatry*, *68*(12), 1257-1266.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*(6), 464-476.

Supplementary Results

Index of Supplementary Results

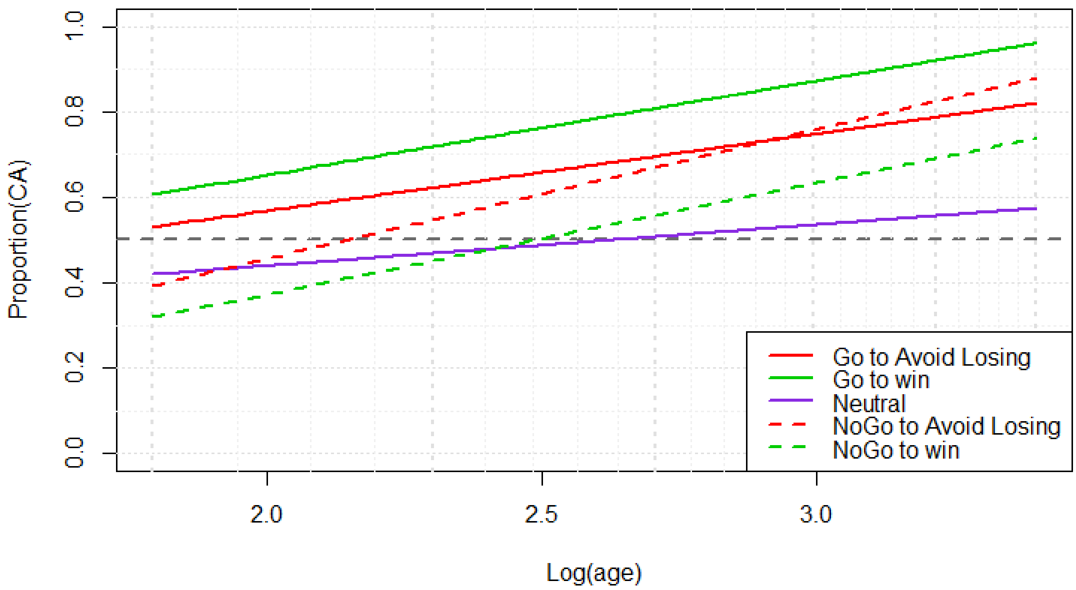
S1. Characterization of each subjects' group (without the 6 excluded participants): its age-range mean, the standard deviation (SD), both male and female frequencies (%), and the absolute frequency of participants. The excluded participants belonged to the groups highlighted in bold.

Subjects' groups	Subjects' age-range	Age-range (mean)	Age-range (SD)	Male (%)	Female (%)	Participants (N)
6	[6.0-7.0)	6.67	0.24	64.3	35.7	14
7	**[7.0 - 7.5)**	**7.20**	**0.11**	**26.7**	**73.3**	**15**
8	[7.5 - 8.5)	8.07	0.30	55.0	45.0	20
9	[8.5 - 9.5)	9.07	0.27	54.6	45.5	22
10	[9.5 - 10.5)	9.86	0.22	55.5	44.4	18
11	[10.5 - 11.5)	11.01	0.28	55.6	44.4	18
12	**[11.5 - 12.5)**	**11.95**	**0.28**	**61.1**	**38.9**	**18**
13	[12.5 - 13.5)	12.99	0.29	47.8	52.2	23
14	**[13.5 - 14.5)**	**14.04**	**0.27**	**29.4**	**70.6**	**17**
15	**[14.5 - 15.5)**	**15.05**	**0.30**	**48.7**	**51.3**	**37**
16	**[15.5 - 16.5)**	**16.01**	**0.31**	**50.0**	**50.0**	**40**
17	[16.5 - 17.5)	17.01	0.28	63.2	36.8	38
18	[17.5 - 18.5)	18.01	0.27	68.6	31.4	35
19-20	[18.5 - 20.5)	19.16	0.80	52.0	48.0	25
21-23	[20.5 - 24.0)	22.47	1.22	45.5	54.6	22
24-29	[24.0 - 30.0)	25.24	1.69	50.0	50.0	20
30-49	**[30.0 - 50.0)**	**39.91**	**6.65**	**33.3**	**66.7**	**18**
49-80	[50.0 - 80.0)	60.86	10.96	38.5	61.5	13
Total		**17.34**	**10.74**	**51.6**	**48.4**	**413**

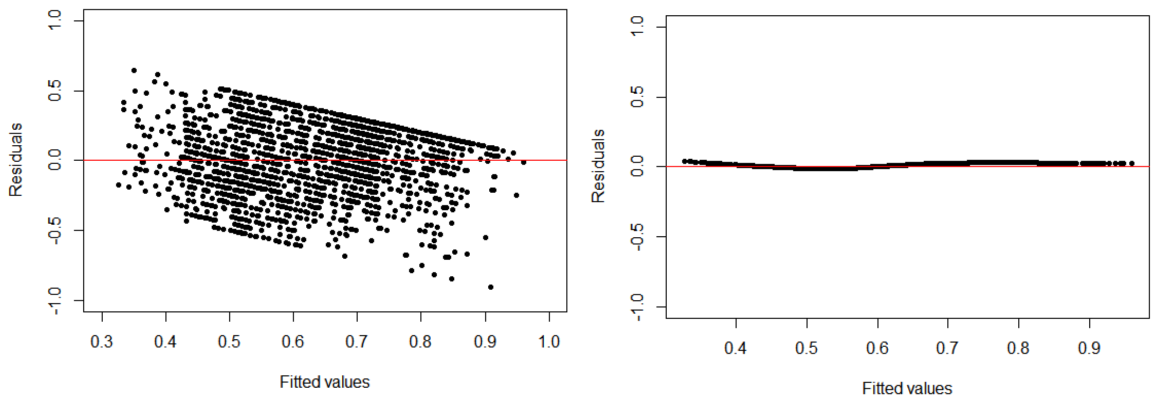
S2. Equations of each task condition of the Model_7.

Conditions	**Model equations**
Go to avoid losing	$p \sim \beta_0 + \beta_2 action + \beta_3 \log(age) + \beta_6 action \text{ x} \log(age)$
Go to win	$p \sim \beta_0 + \beta_2 action + \beta_3 \log(age) + \beta_6 action \text{ x} \log(age) + \beta_7 valence \text{ x } action \text{ x} \log(age)$
Neutral	$p \sim \beta_0 + \beta_1 neutral + \beta_3 \log(age) + \beta_5 neutral \text{ x} \log(age)$
NoGo to avoid losing	$p \sim \beta_0 + \beta_3 \log(age)$
NoGo to win	$p \sim \beta_0 + \beta_3 \log(age) + \beta_4 valence \text{ x} \log(age)$

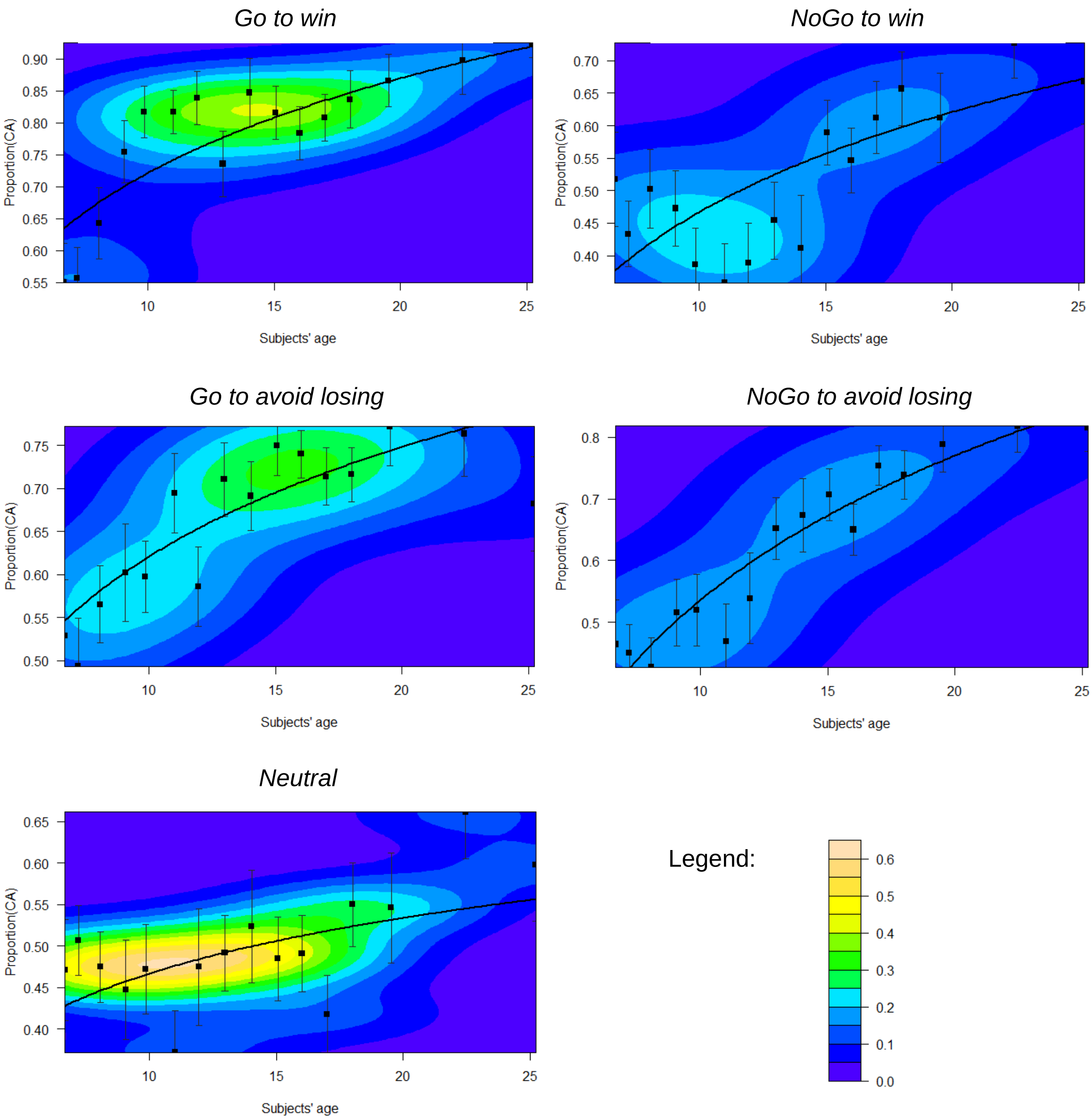
S3. The predicted learning curves of the model_7 for all task conditions. Each straight line represents one task condition as indicated by the legend.



S4. The residuals versus the fitted values of the model_7: all conditions. Left - residuals versus the fitted values of all conditions. Right – results of the LOWESS function applied to the residuals versus the fitted values of all conditions.



S5. Heat-maps of the kernel density estimation of the Proportion(CA) for each task condition, with the prediction learning curves of the model_7 (black lines), and the averages of the proportion(CA) for each of the 16 subjects' groups (black points). From left to right and top to bottom: *Go to win*, *NoGo to win*, *Go to avoid losing*, *NoGo to avoid losing*, and *Neutral* conditions; and heat-maps' legend.



S6. The contrasts weights and their respective values of both intercept and slope of each task condition.

	Conditions	Contrasts weights	Values
Intercept	*Go to avoid losing*	$\beta_0 + \beta_2$	0.210
	Go to win	$\beta_0 + \beta_2$	0.210
	Neutral	$\beta_0 + \beta_1$	0.248
	NoGo to avoid losing	β_0	-0.148
	NoGo to win	β_0	-0.148
Slope	*Go to avoid losing*	$\beta_3 + \beta_6$	0.179
	Go to win	$\beta_3 + \beta_4 + \beta_6 + \beta_7$	0.221
	Neutral	$\beta_3 + \beta_5$	0.096
	NoGo to avoid losing	β_3	0.302
	NoGo to win	$\beta_3 + \beta_4$	0.260

S7. Relative frequencies of the subjects' answers concerning their attention during the task performance.

	Subjects' self-evaluation about their task performance	
Subjects' Group	**# With Attention**	**# Without Attention**
6	0.64	0.36
7	0.53	0.47
8	0.70	0.30
9	0.82	0.18
10	0.83	0.17
11	0.78	0.22
12	0.83	0.17
13	0.74	0.26
14	0.82	0.18
15	0.65	0.35
16	0.78	0.23
17	0.61	0.39
18	0.83	0.17
19-20	0.64	0.36
21-23	0.77	0.23
24-29	0.75	0.25
30-49	0.67	0.33
50-80	0.77	0.23
Total	0.73	0.27

S8. Relative frequencies of the subjects' answers concerning which part(s) of the task they didn't pay attention to their task performance: 1st, 2nd, 3rd blocks and/or during the entire task.

Subjects' Group	1st Block	2nd Block	3rd Block	Entire Task
6	0.20	0	0.80	0
7	0.14	0.36	0.36	0.14
8	0	0.50	0.50	0
9	0.25	0.63	0.13	0
10	0.67	0	0	0.33
11	0.25	0.63	0.13	0
12	0.67	0	0.33	0
13	0.17	0.33	0.33	0.17
14	0.33	0	0.67	0
15	0.31	0.15	0.31	0.23
16	0	0.28	0.61	0.11
17	0.13	0.27	0.40	0.20
18	0.17	0.25	0.42	0.17
19-20	0.22	0.22	0.44	0.11
21-23	0.30	0.50	0.20	0
24-29	0.70	0.10	0.20	0
30-49	0.25	0.42	0.33	0
50-80	0	0.33	0.33	0.33
Total	0.23	0.28	**0.38**	0.12

S9. Absolute frequency of the subjects' answers of the amount of different task's images

Subjects' answers	Absolute frequency
0	1
1	2
2	1
3	6
4	24
5	183
6	93
7	39
8	17
9	11
10	18
11	1
12	6
13	1
15	1
16	1
20	3
26	1
27	1
30	1
70	1
1 e+18	1

Appendices

Index of Appendices

A1. Consent form delivered to youngest participants

Estudo de Investigação

O estudo "Neuropsicologia de Perturbações Psiquiátricas do Desenvolvimento", coordenado pelo Professor Doutor Tiago Vaz Maia do Instituto de Medicina Molecular/Faculdade de Medicina da Universidade de Lisboa, tenta perceber o que se passa no cérebro de crianças e adolescentes com doenças psiquiátricas. Para fazer este estudo, os investigadores precisam da participação de crianças e adolescentes com doenças psiquiátricas, mas também precisam da participação de crianças e adolescentes saudáveis.

Parte do estudo consiste em fazer alguns jogos de computador que os investigadores desenvolveram. Numa primeira fase, os investigadores precisam que crianças e adolescentes saudáveis joguem estes jogos, para os testar e adequar o seu grau de dificuldade às várias idades. É para esta fase que estamos a pedir a tua participação.

Se aceitares participar, vais jogar estes jogos. Além disso, poderás também fazer outras provas, se aceitares, que são:

(Investigador: Assinalar o que se aplica.)

- ☐ Usar um aparelho muito pequeno (preso no cinto das calças, à volta do pulso como um relógio ou noutro sítio conveniente) que grava a quantidade de movimento que fazes;
- ☐ Gravar os movimentos da tua pupila enquanto estás a fazer os jogos em computador.

A participação neste estudo é importante para ajudar a compreender melhor algumas doenças psiquiátricas em crianças e adolescentes.

A participação é totalmente voluntária e pode ser interrompida a qualquer momento do estudo. A identificação dos participantes nunca será divulgada, por isso só os investigadores poderão saber quem são os participantes do estudo.

Declaração de Consentimento

Declaro que aceito participar no estudo "Neuropsicologia de Perturbações Psiquiátricas do Desenvolvimento", coordenado pelo Professor Doutor Tiago Vaz Maia do Instituto de Medicina Molecular/Faculdade de Medicina da Universidade de Lisboa.

Fui informado que irei realizar alguns jogos em computador.

Aceito também:

(Investigador: Riscar o que não se aplica. Participante: Colocar uma cruz conforme o caso: A – Aceito; NA – Não Aceito)

	A	NA
1 – Utilizar um aparelho para analisar os meus movimentos		
2 – Gravar os movimentos da minha pupila enquanto faço os jogos em computador		

Percebo que a minha participação é voluntária – isto é, só participo neste estudo se eu quiser – e que se decidir não participar isso não terá qualquer consequência negativa para mim. Para além disso, posso decidir parar a minha participação no estudo a qualquer altura, ou recusar participar em qualquer componente do estudo que eu não queira fazer.

Compreendo que os resultados deste estudo poderão ser publicados em revistas científicas ou apresentados em conferências científicas, mas sempre sem o meu nome ou qualquer outra informação que pudesse servir para saberem quem eu sou. Só os investigadores directamente envolvidos neste estudo saberão quem eu sou.

Se havia coisas que eu não percebi bem, perguntei ao investigador que me está a explicar o estudo, e ele/a esclareceu todas as minhas dúvidas.

Data ___/___/___

Nome completo do participante:____

Assinatura do participante: ____

Nome do investigador: ____

Assinatura do investigador:____

Estudo de Investigação

O estudo "Neuropsicologia de Perturbações Psiquiátricas do Desenvolvimento", coordenado pelo Professor Doutor Tiago Vaz Maia do Instituto de Medicina Molecular/Faculdade de Medicina da Universidade de Lisboa, tem por objectivo investigar dificuldades cognitivas em Perturbações do Desenvolvimento, tais como Perturbação de Hiperactividade com Défice de Atenção, Síndrome de Tourette, Perturbação Obsessivo-Compulsiva, entre outras. Os participantes deste estudo serão crianças e adolescentes.

Neste projecto serão aplicadas algumas tarefas cognitivas na forma de jogos de computador simples, desenvolvidos pelos investigadores do projecto, bem como algumas técnicas novas. Numa fase inicial, é necessário realizar testes piloto com crianças e adolescentes saudáveis. Esta fase do estudo irá permitir a optimização as tarefas em termos analíticos e técnicos, como, por exemplo, garantir que a dificuldade das tarefas é apropriadas para a idade dos participantes do estudo e que as mesmas permitem observar as características que desejamos. É para esta fase que estamos a pedir a participação do menor de quem é representante legal.

A participação nesta fase do estudo envolve realizar algumas tarefas neuropsicológicas em computador.

Outros procedimentos poderão também ser realizados, nomeadamente:
(Investigador: Assinalar o que se aplica.)

- ☐ Análise da atividade motora do participante usando um acelerómetro. Os acelerómetros são dispositivos muito pequenos que podem ser presos na roupa ou usados como um relógio de pulso, por isso o seu uso não é incómodo.

- ☐ Observação de alterações no diâmetro da pupila enquanto o participante realiza as tarefas em computador. O registo será feito através de aparelhos semelhantes a pequenas câmaras de vídeo. É um procedimento simples e não invasivo.

Caso os dados decorrentes da execução das tarefas sejam apresentados publicamente em eventos académicos e científicos, será sempre garantida a sua confidencialidade de modo a que a identidade de cada criança/adolescente não seja revelada.

A participação neste estudo é inteiramente voluntária e poderá ser interrompida a qualquer momento, sem prejuízo para o participante ou para o representante legal do mesmo.

Declaração de Consentimento

Declaro que aceito que o menor ______________________________ ______________________________, de quem sou representante legal, participe na fase de testes piloto do estudo "Neuropsicologia de Perturbações Psiquiátricas do Desenvolvimento", coordenado pelo Professor Doutor Tiago Vaz Maia do Instituto de Medicina Molecular/Faculdade de Medicina da Universidade de Lisboa.

Fui informado que o menor, que represento, irá realizar algumas tarefas neuropsicológicas computorizadas.

Para além disso, autorizo a participação do menor que represento nos seguintes pontos abaixo discriminados:

(Investigador: Riscar o que não se aplica. Representante legal: Colocar uma cruz conforme o caso: A – Autorizo; NA – Não Autorizo)

	A	NA
1 – Utilização de acelerómetros para análise da atividade motora		
2 – Análise de variações do diâmetro da pupila durante as tarefas em computador		

Foi-me também indicado que a identidade do menor, que represento, será sempre salvaguardada e nunca divulgada publicamente.

Compreendo que a participação do menor acima mencionado neste estudo é inteiramente voluntária e que se eu ou ele/a não quisermos que ele/a participe no estudo daí não advirá qualquer consequência negativa para nós. Para além disso, poderemos parar a participação nos estudo a qualquer altura, ou recusar participar em qualquer componente do estudo, sem qualquer tipo de prejuízo.

Este estudo mereceu o parecer favorável da Comissão de Ética do Centro Hospitalar Lisboa Norte / Faculdade de Medicina da Universidade de Lisboa.

Declaro ter lido e compreendido este documento e que o/a investigador/a abaixo assinado me esclareceu quaisquer dúvidas que eu pudesse ter.

Data ___/___/___

Nome completo do participante:____________________________________

Nome completo do representante legal:______________________________

Assinatura do representante legal: _________________________________

Nome do investigador:____

Assinatura do investigador:______________________________________

A2. Consent form delivered to adult participants

Estudo de Investigação

O estudo "Neuropsicologia de Perturbações Psiquiátricas do Desenvolvimento", coordenado pelo Professor Doutor Tiago Vaz Maia do Instituto de Medicina Molecular/Faculdade de Medicina da Universidade de Lisboa, tem por objectivo investigar dificuldades cognitivas em Perturbações do Desenvolvimento, tais como Perturbação de Hiperactividade com Défice de Atenção, Síndrome de Tourette, Perturbação Obsessivo-Compulsiva, entre outras. Os participantes deste estudo serão adultos.

Neste projecto serão aplicadas algumas tarefas cognitivas na forma de jogos de computador simples, desenvolvidos pelos investigadores do projecto, bem como algumas técnicas novas. Numa fase inicial, é necessário realizar testes piloto com adultos saudáveis. Esta fase do estudo irá permitir a optimização das tarefas em termos analíticos e técnicos, como, por exemplo, garantir que a dificuldade das tarefas é apropriadas para a idade dos participantes do estudo e que as mesmas permitem observar as características que desejamos. Estamos a pedir a sua participação para esta fase.

A participação nesta fase do estudo envolve realizar algumas tarefas neuropsicológicas em computador. Outros procedimentos poderão também ser realizados, nomeadamente:

(Investigador: Assinalar o que se aplica.)

- ☐ Análise da atividade motora do participante usando um acelerómetro. Os acelerómetros são dispositivos muito pequenos que podem ser presos na roupa ou usados como um relógio de pulso, por isso o seu uso não é incómodo.

- ☐ Observação de alterações no diâmetro da pupila enquanto o participante realiza as tarefas em computador. O registo será feito através de aparelhos semelhantes a pequenas câmaras de vídeo. É um procedimento simples e não invasivo.

Caso os dados decorrentes da execução das tarefas sejam apresentados publicamente em eventos académicos e científicos, será sempre garantida a sua confidencialidade de modo a que a identidade de cada participante não seja revelada.

Pode, em qualquer altura e sem qualquer prejuízo, interromper a sua participação no estudo. Este estudo mereceu parecer favorável da Comissão de Ética do Centro Hospitalar Lisboa Norte / Faculdade de Medicina da Universidade de Lisboa.

Declaração de Consentimento

Eu, ____,
declaro que aceito participar na fase de testes piloto do estudo "Neuropsicologia de Perturbações Psiquiátricas do Desenvolvimento", coordenado pelo Professor Doutor Tiago Vaz Maia do Instituto de Medicina Molecular/Faculdade de Medicina da Universidade de Lisboa.

Fui informado que de que esta fase do estudo envolve realizar algumas tarefas neuropsicológicas computorizadas.

Para além disso, aceito também a minha participação nos seguintes pontos abaixo discriminados:

(Investigador: Riscar o que não se aplica. Participante: Colocar uma cruz conforme o caso: A – Autorizo; NA – Não Autorizo)

	A	NA
1 – Utilização de acelerómetros para análise da atividade motora		
2 – Análise de variações do diâmetro da pupila durante as tarefas em computador		

Foi-me também indicado que a minha identidade será sempre salvaguardada e nunca divulgada publicamente.

Compreendo que a minha participação neste estudo é inteiramente voluntária e que se eu não quiser participar no estudo daí não advirá qualquer consequência negativa para mim. Para além disso, sei que posso parar a minha participação no estudo a qualquer altura, ou recusar participar em qualquer componente do estudo, sem qualquer tipo de prejuízo para mim.

Compreendo que os resultados deste estudo poderão vir a ser publicados em revistas científicas ou apresentados em conferências científicas, sendo, no entanto, garantida a confidencialidade de todos os dados. A minha identidade nunca será relevada em qualquer descrição ou publicação deste trabalho.

Este estudo mereceu o parecer favorável da Comissão de Ética do Centro Hospitalar Lisboa Norte / Faculdade de Medicina da Universidade de Lisboa.

Declaro ter lido e compreendido este documento e que o/a investigador/a abaixo assinado me esclareceu quaisquer dúvidas que eu pudesse ter.

Data ___/___/___

Nome completo do participante:____

Assinatura participante: ____

Nome do investigador:____

Assinatura do investigador:____

A3. Instruction on screen (displayed by 3 screens)

INSTRUÇÕES

- Vão aparecer várias imagens ao longo do jogo.
- Cada vez que aparecer uma imagem, tens de escolher carregar ou não na **barra de espaços**. Se não sabes onde é a barra de espaços, pergunta ao instrutor.
- A seguir a carregares ou não carregares, aparece o resultado da tua escolha. Podes:
    - ganhar 1 ponto (+1)
    - perder 1 ponto (-1)
    - não ganhar nem perder pontos (0)
- O objetivo do jogo é conseguires ter o maior número de pontos possíveis.
- Para algumas imagens, o resultado costuma ser melhor se carregares na barra de espaços; para outras imagens, o resultado costuma ser melhor se não carregares; para outras, pode não fazer diferença se carregas ou não carregas. No início, tu não vais saber para quais imagens é melhor carregar e para quais é melhor não carregar, mas como cada imagem vai aparecer várias vezes, podes conseguir aprender para quais imagens deves carregar e para quais não deves.
- Para a mesma imagem, a mesma ação (carregar ou não carregar) pode dar-te resultados diferentes. Isso é normal, porque neste jogo as ações não dão sempre o mesmo resultado. Mesmo assim, há imagens para as quais é melhor carregar porque regra geral isso dá melhor resultado, e há outras imagens para as quais é melhor não carregar porque regra geral isso dá melhor resultado.
- As regras do jogo não mudam ao longo do jogo. As imagens para as quais é melhor carregar e não carregar são as mesmas ao longo do jogo. O jogo tem três partes, mas não muda nada entre essas partes. Os intervalos são só para descansares um bocadinho.
- No início do jogo deves experimentar carregar e não carregar na tecla para todas as imagens para descobrires qual é a melhor situação em cada imagem.
- Quando decidires carregar na tecla tenta carregar o mais rápido possível.
    - Antes de continuares, chama o instrutor.

A4. Additional instruction (to read)

- Okay, vou-te explicar brevemente as instruções. Se tiveres alguma dúvida acerca do que leste ou acerca do que eu explicar, pergunta!
- Vai aparecer uma imagem de cada vez no ecrã. As imagens são deste género (mostrar uma imagem no ecrã, que deverá estar num ecrã depois das instruções que o sujeito leu).
- De cada vez que aparece uma imagem deste género, tu vais escolher carregar ou não nesta tecla (indicar qual é a tecla).
- A seguir, aparece o resultado daquilo que fizeste: aparece escrito no ecrã se ganhaste um ponto (+1), se perdeste um ponto (-1), ou se não ganhaste nem perdeste (0). Mesmo quando não carregas na tecla vai aparecer-te um resultado. Os pontos que aparecerem contam para a tua pontuação final, quer tenhas carregado na tecla ou não.
- Para a mesma imagem, a mesma ação (carregar ou não carregar) pode dar resultados diferentes. Por exemplo, esta imagem não vai mesmo aparecer no jogo (apontar para imagem no ecrã), mas faz de conta que aparecia. Podia acontecer a imagem aparecer e tu carregares na tecla e ganhares um ponto. Depois apareciam outras imagens e quando voltasse a aparecer esta imagem tu podias voltar a carregar mas perder um ponto ou teres zero. Mesmo assim, podia ser melhor carregares de cada vez que aparecesse esta imagem porque podia fazer-te ganhar pontos mais vezes do que perder. Também podia ser ao contrário, fazer-te perder pontos mais vezes do que ganhar. Por isso tens de experimentar umas quantas vezes para cada imagem para saberes o que é melhor fazer.
- As regras do jogo são muito simples: não há sequências nem alterações ao longo do jogo. Só tens de aprender para que imagens é melhor carregares e para que imagens é melhor não carregares.
- Tens alguma pergunta? (Responder a todas as perguntas.)
- Okay, então agora vou-te mostrar como o jogo funciona.

A5. Final questionnaire

Jogo

Olá! Espero que te tenhas divertido! =)

Por favor, responde às seguintes questões. Se tiveres dúvidas pergunta ao instrutor.

Mais uma vez, obrigada pela tua participação!

*Obrigatório

1. **Número de Identificação do participante** *
   (Chama o instrutor)

   ____

Informações Pessoais

2. **Data de realização do jogo** *

   ____
   Exemplo: 15 de dezembro 2012

3. **Idade** *
   (Coloca apenas o número correspondente à tua idade no local da resposta)

   ____

4. **Data de Nascimento** *

   ____
   Exemplo: 15 de dezembro 2012

5. **Sexo** *
   Marcar apenas uma oval.
   - ◯ Feminino
   - ◯ Masculino

6. **Situação profissional** *
   Marcar apenas uma oval.
   - ◯ Estudante ou Trabalhador/Estudante *Passe para a pergunta 10.*
   - ◯ Outro *Passe para a pergunta 7.*

Outra situação profissional

7. **Em que situação profissional te encontras** *
   Marcar apenas uma oval.
   - ◯ Empregada/o
   - ◯ Desempregada/o
   - ◯ Reformada/o

Habilitações Literárias

8. **Qual a última habilitação literária obtida?** *
   Marcar apenas uma oval.
   - ◯ 4º ano *Passe para a pergunta 9.*
   - ◯ 7º ano *Passe para a pergunta 9.*
   - ◯ 9º ano *Passe para a pergunta 9.*
   - ◯ 10º ano *Passe para a pergunta 9.*
   - ◯ 11º ano *Passe para a pergunta 9.*
   - ◯ 12º ano *Passe para a pergunta 9.*
   - ◯ Licenciatura/Bacharelato *Passe para a pergunta 9.*
   - ◯ Mestrado *Passe para a pergunta 9.*
   - ◯ Doutoramento *Passe para a pergunta 9.*
   - ◯ Pós-Graduação *Passe para a pergunta 9.*
   - ◯ Outra: ____________________ *Passe para a pergunta 9.*

Área Profissional

9. **Área profissional exercida** *
   Marcar apenas uma oval.
   - ◯ Ciências e Tecnologias
   - ◯ Ciências da Saúde
   - ◯ Ciências Sociais e Humanas
   - ◯ Artes ou Desporto
   - ◯ Outra: ____________________

Estudante

10. **Ano de escolaridade** *
(Escolhe o ano de escolaridade que atualmente frequentas)
Marcar apenas uma oval.

- 1º ano *Passe para a pergunta 11.*
- 2º ano *Passe para a pergunta 11.*
- 3º ano *Passe para a pergunta 11.*
- 4º ano *Passe para a pergunta 11.*
- 5º ano *Passe para a pergunta 11.*
- 6º ano *Passe para a pergunta 11.*
- 7º ano *Passe para a pergunta 11.*
- 8º ano *Passe para a pergunta 11.*
- 9º ano *Passe para a pergunta 11.*
- 10º ano *Passe para a pergunta 12.*
- 11º ano *Passe para a pergunta 12.*
- 12º ano *Passe para a pergunta 12.*
- Ensino Superior *Passe para a pergunta 15.*

Ensino Básico e Secundário

11. **Qual a escola que frequentas atualmente?** *
Marcar apenas uma oval.

- Escola EB1/J1 Miratejo *Passe para a pergunta 18.*
- Escola Básica 2, 3 de Corroios *Passe para a pergunta 18.*
- Escola Secundária João de Barros *Passe para a pergunta 18.*
- Outra: ________________ *Passe para a pergunta 18.*

Ensino Secundário

12. **Seleciona o curso que frequentas atualmente** *
Marcar apenas uma oval.

- Curso científico Humanístico *Passe para a pergunta 13.*
- Curso Profissional *Passe para a pergunta 14.*
- CEF *Passe para a pergunta 11.*

Cursos Científicos e Humanísticos

13. **Seleciona o curso que frequentas atualmente** *
Marcar apenas uma oval.

- Ciências e Tecnologias *Passe para a pergunta 11.*
- Línguas e Humanidades *Passe para a pergunta 11.*
- Ciências Socioeconómicas *Passe para a pergunta 11.*
- Artes Visuais *Passe para a pergunta 11.*

Cursos Profissionais

14. **Seleciona o curso que frequentas atualmente** *
Marcar apenas uma oval.

- Gestão e Manutenção de Equipamentos Informáticos *Passe para a pergunta 11.*
- Gestão e Programação de Sistemas Informáticos *Passe para a pergunta 11.*

Ensino Superior

15. **Seleciona o que frequentas atualmente** *
Marcar apenas uma oval.

- Licenciatura / Mestrado integrado 1º ciclo
- Mestrado / Mestrado integrado 2º ciclo
- Doutoramento
- Pós-Graduação

16. **Seleciona o estabelecimento de ensino superior que frequentas atualmente?** *
Na opção "Outro" escreve o estabelecimento de ensino superior de igual forma aos exemplos apresentados: sigla da faculdade ou escola superior, hífen, sigla da universidade ou politécnico (quando aplicado).
Marcar apenas uma oval.

- FCT-UNL
- FM-UL
- IST-UL
- Outra: ________________

17. **Seleciona a tua área de estudo principal** *
(Se tiveres dúvidas pergunta ao instrutor)
Marcar apenas uma oval.

- Ciências e Tecnologias
- Ciências da Saúde
- Ciências Sociais e Humanas
- Artes e Desporto
- Outra: ....

As próximas questões são acerca do jogo que acabaste de jogar.

Atenção durante o jogo

18. **Achas que estiveste sempre atento durante o jogo?** *
Marcar apenas uma oval.

- Sim, estive sempre atento durante todo o jogo. *Passe para a pergunta 20.*
- Não, nem sempre consegui estar atento durante todo o jogo.

19. **Em que parte(s) do jogo não estiveste atento?** *
(Podes escolher mais do que uma opção)
Marcar tudo o que for aplicável.

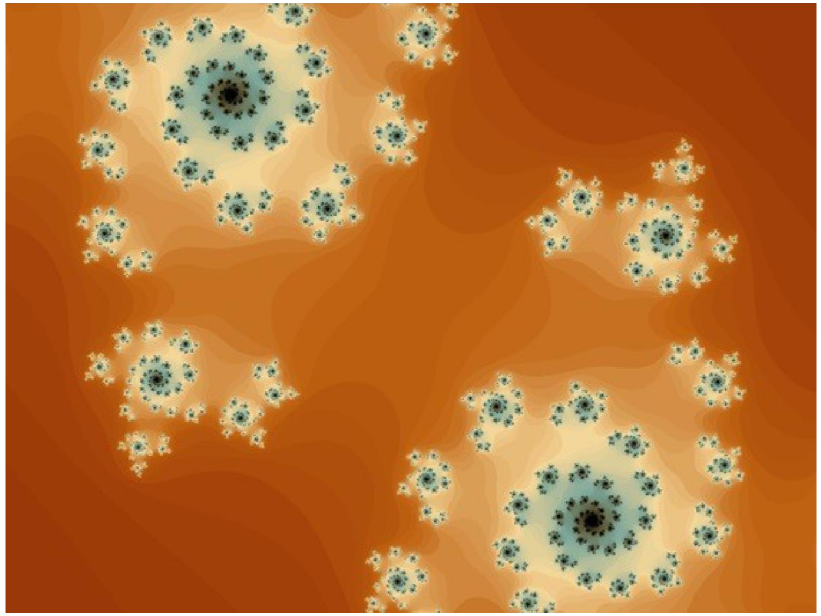
- [ ] 1ª Parte.
- [ ] 2ª Parte.
- [ ] 3ª Parte.
- [ ] Todo o jogo.

20. **Quantas imagens diferentes achas que o jogo tinha?** *
(Coloca apenas o número no local da resposta)

....

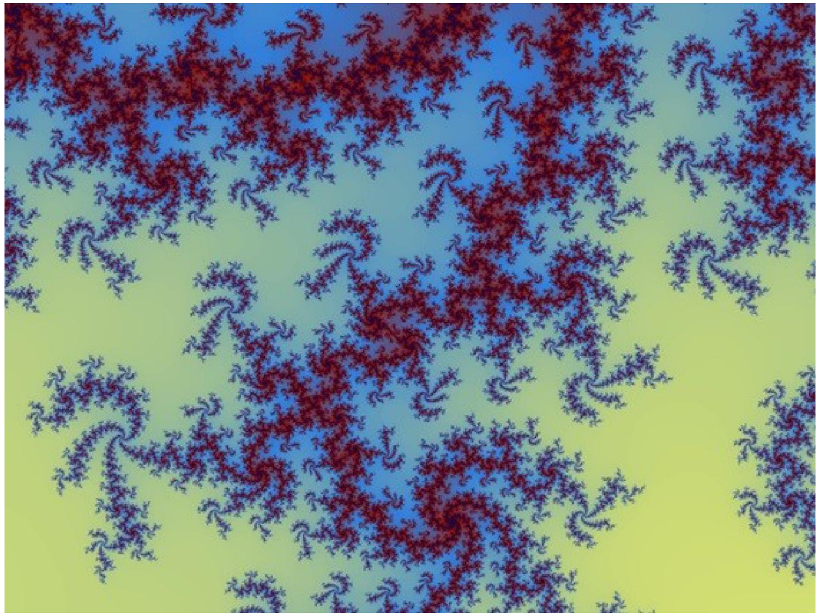
Beleza das imagens

As próximas questões são apenas referentes à beleza de cada imagem. Não tem nada a ver com o que fizeste no jogo.



21. **Como classificas esta imagem quanto à sua beleza?** *
(Quanto achas a imagem bonita?)
Marcar apenas uma oval por linha.

	Muito feia	Mais ao menos feia	Nem feia nem bonita	Mais ao menos bonita	Muito bonita
	◯	◯	◯	◯	◯

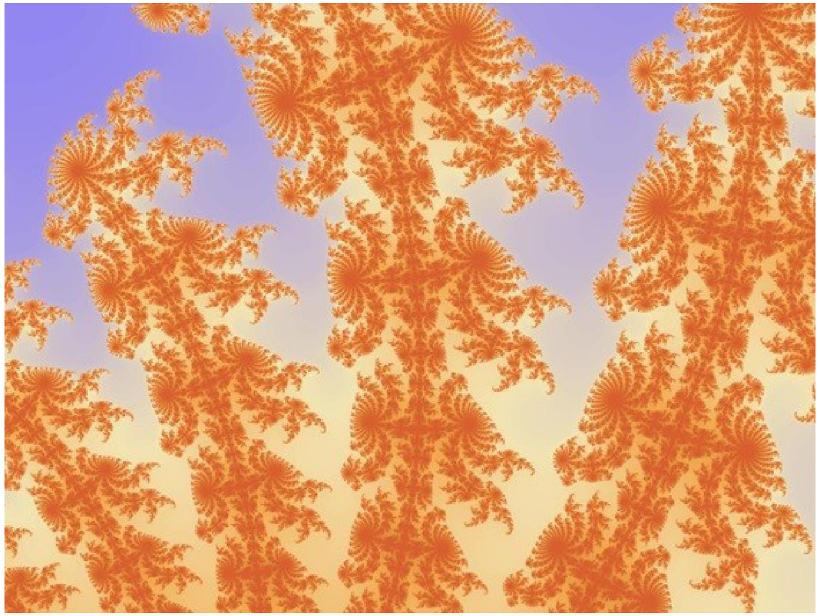


22. **Como classificas esta imagem quanto à sua beleza?** *

(Quanto achas a imagem bonita?)

Marcar apenas uma oval por linha.

	Muito feia	Mais ao menos feia	Nem feia nem bonita	Mais ao menos bonita	Muito bonita
	◯	◯	◯	◯	◯

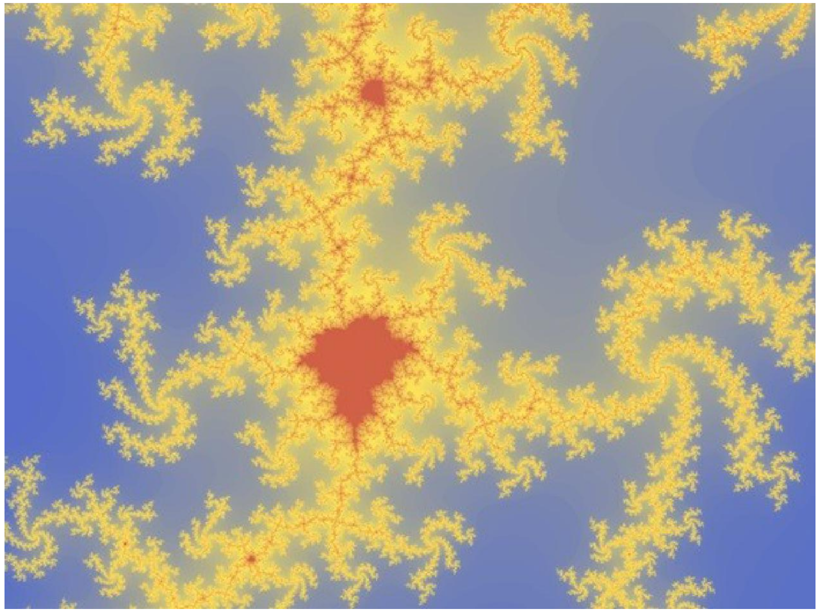


23. **Como classificas esta imagem quanto à sua beleza?** *

(Quanto achas a imagem bonita?)

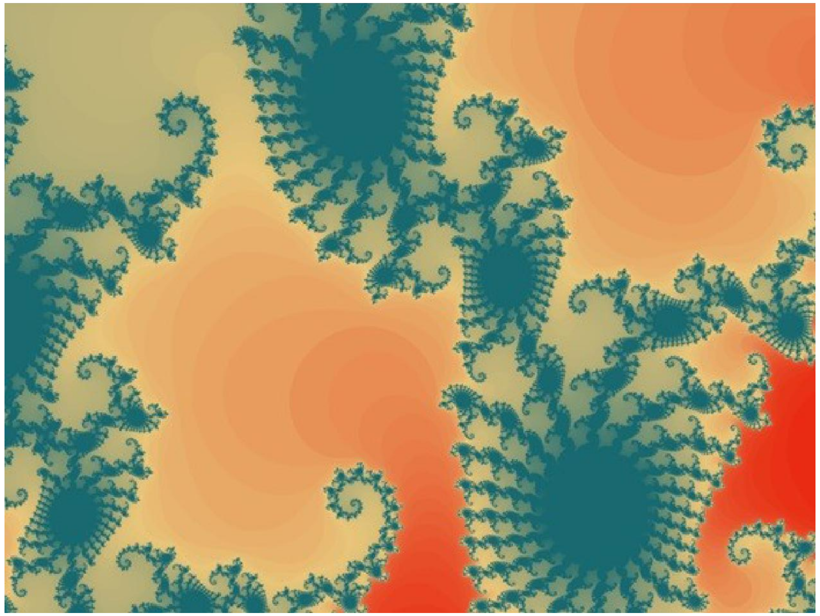
Marcar apenas uma oval por linha.

	Muito feia	Mais ao menos feia	Nem feia nem bonita	Mais ao menos bonita	Muito bonita
	◯	◯	◯	◯	◯



24. **Como classificas esta imagem quanto à sua beleza?** *
(Quanto achas a imagem bonita?)
Marcar apenas uma oval por linha.

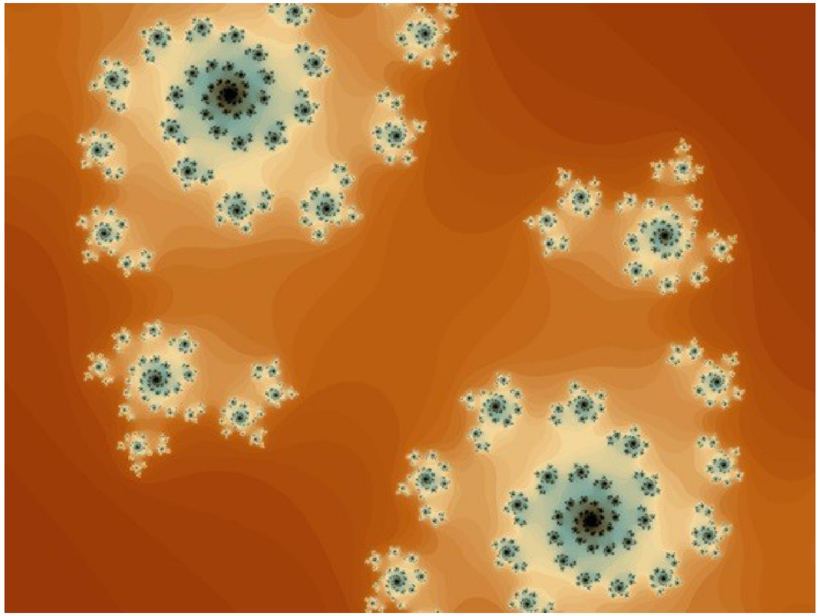
	Muito feia	Mais ao menos feia	Nem feia nem bonita	Mais ao menos bonita	Muito bonita
	◯	◯	◯	◯	◯



25. **Como classificas esta imagem quanto à sua beleza?** *
(Quanto achas a imagem bonita?)
Marcar apenas uma oval por linha.

	Muito feia	Mais ao menos feia	Nem feia nem bonita	Mais ao menos bonita	Muito bonita
	◯	◯	◯	◯	◯

As próximas questões são referentes ao que fizeste no jogo. É normal que demores algum tempo para te recordares daquilo que fizeste.



26. **Nesta imagem, o que era melhor fazeres?** *
Marcar apenas uma oval.

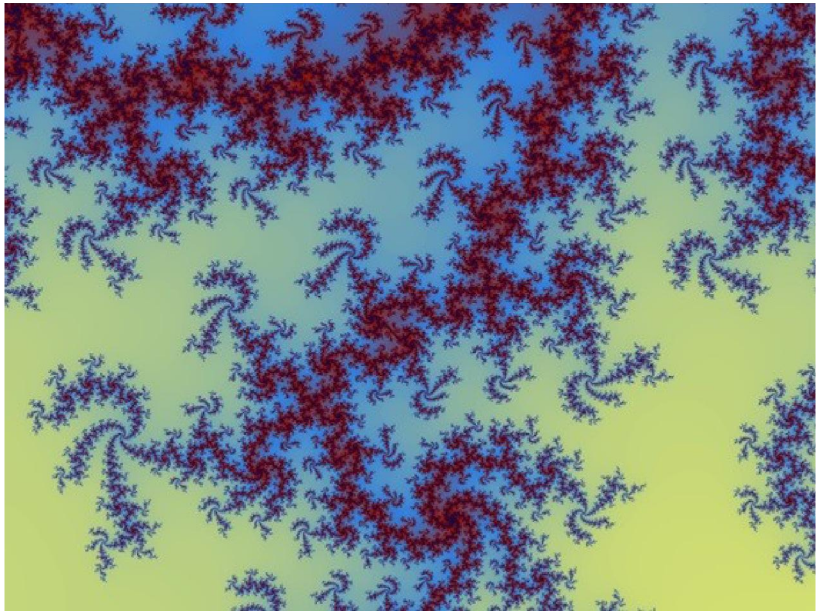
- Carregar na barra de espaços.
- Não carregar na barra de espaços.
- Tanto fazia carregar ou não na barra de espaços.
- Não sei / Não me lembro.

27. **Se CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- -1
- 0
- +1
- Não sei / Não me lembro

28. **Se NÃO CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- -1
- 0
- +1
- Não sei / Não me lembro



29. **Nesta imagem, o que era melhor fazeres?** *
Marcar apenas uma oval.

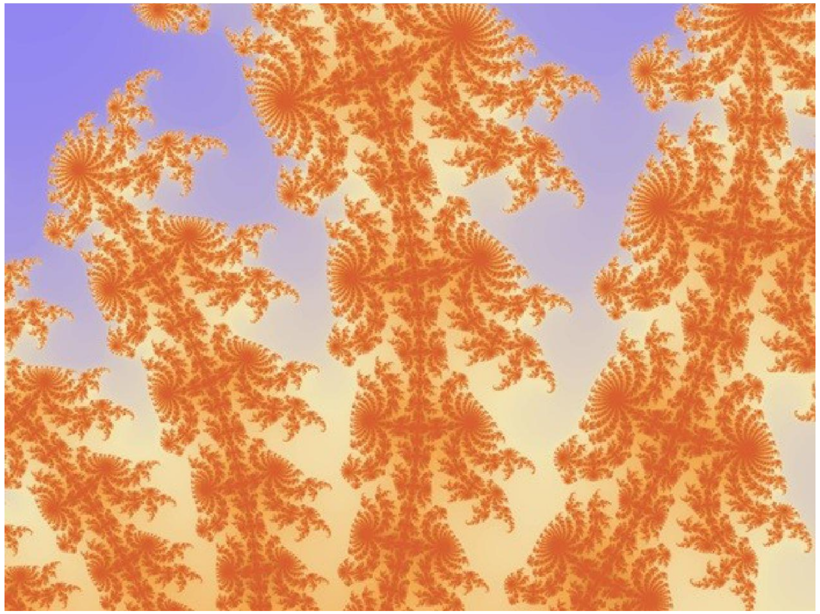
- Carregar na barra de espaços.
- Não carregar na barra de espaços.
- Tanto fazia carregar ou não na barra de espaços.
- Não sei / Não me lembro.

30. **Se CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- ( ) -1
- ( ) 0
- ( ) +1
- ( ) Não sei / Não me lembro

31. **Se NÃO CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- ( ) -1
- ( ) 0
- ( ) +1
- ( ) Não sei / Não me lembro



32. **Nesta imagem, o que era melhor fazeres?** *
Marcar apenas uma oval.

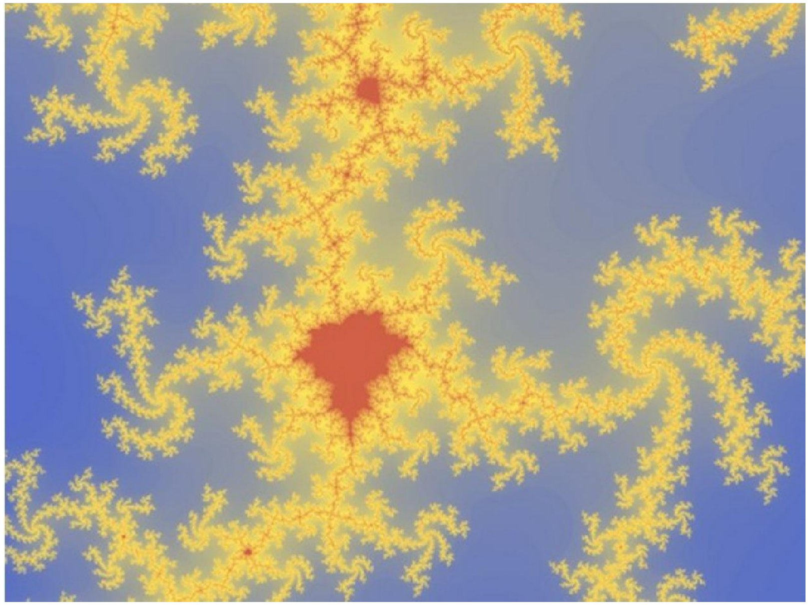
- ( ) Carregar na barra de espaços.
- ( ) Não carregar na barra de espaços.
- ( ) Tanto fazia carregar ou não na barra de espaços.
- ( ) Não sei / Não me lembro.

33. **Se CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- ( ) -1
- ( ) 0
- ( ) +1
- ( ) Não sei / Não me lembro

34. **Se NÃO CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- ( ) -1
- ( ) 0
- ( ) +1
- ( ) Não sei / Não me lembro



35. **Nesta imagem, o que era melhor fazeres?** *
Marcar apenas uma oval.

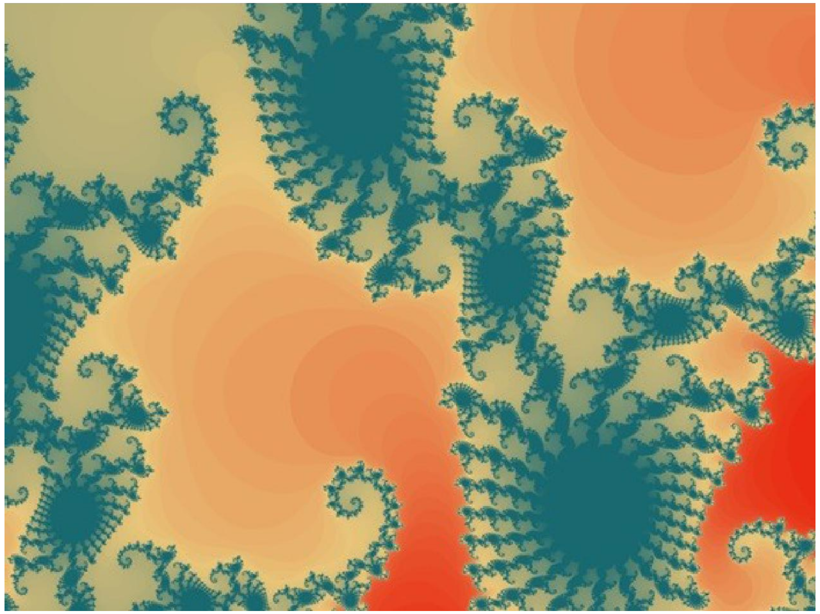
- ( ) Carregar na barra de espaços.
- ( ) Não carregar na barra de espaços.
- ( ) Tanto fazia carregar ou não na barra de espaços.
- ( ) Não sei / Não me lembro.

36. **Se CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- ( ) -1
- ( ) 0
- ( ) +1
- ( ) Não sei / Não me lembro

37. **Se NÃO CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- ( ) -1
- ( ) 0
- ( ) +1
- ( ) Não sei / Não me lembro



38. **Nesta imagem, o que era melhor fazeres?** *
Marcar apenas uma oval.

- ( ) Carregar na barra de espaços.
- ( ) Não carregar na barra de espaços.
- ( ) Tanto fazia carregar ou não na barra de espaços.
- ( ) Não sei / Não me lembro.

39. **Se CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- -1
- 0
- +1
- Não sei / Não me lembro

40. **Se NÃO CARREGASSES na barra de espaços quando aparecia esta imagem, que pontos é que costumavam aparecer mais vezes no ecrã?** *
Marcar apenas uma oval.

- -1
- 0
- +1
- Não sei / Não me lembro